GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:08:07 ; Search time 234 Seconds
(without alignments)
1761.178 Million cell updates/sec

Title: US-09-660-568-49
Perfect score: 183
Sequence: 1 cacacactcccccattctga.....ctgtctcctggagctcacca 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	183	100.0	183	19	AAV38517	DNA marker of meta
2	183	100.0	183	20	AAZ30719	Human UC Band #321
3	41	22.4	2113	23	AAS77249	DNA encoding novel
4	35.2	19.2	164	22	ABA51400	Human breast cell
5	35.2	19.2	164	22	ABA69411	Human foetal liver
6	35.2	19.2	164	22	ABA36346	Probe #14812 for g
7	35.2	19.2	164	22	AAK17679	Human brain expres
8	35.2	19.2	164	22	AAK43494	Human bone marrow
9	35.2	19.2	164	22	AAI24280	Probe #14213 for g
10	33.6	18.4	1524	23	ABL30203	Drosophila melanog
c 11	32.8	17.9	3819	23	ABL30202	Drosophila melanog
12	31.2	17.0	3698	22	AAK73760	Human immune/haema
13	31.2	17.0	3698	22	AAK83893	Human immune/haema
14	31.2	17.0	10548	22	AAK64703	Human immune/haema
15	31.2	17.0	10548	22	AAK69813	Human immune/haema
16	31.2	17.0	10548	22	AAK73759	Human immune/haema
17	31.2	17.0	10548	22	AAK80461	Human immune/haema
18	31.2	17.0	10548	22	AAK83894	Human immune/haema
c 19	30.8	16.8	5616	22	AAK79614	Human immune/haema
20	30.6	16.7	3387	23	AAS91234	DNA encoding novel
21	30.6	16.7	5216	22	AAS28261	Genomic sequence #
c 22	30.6	16.7	5216	22	AAS31528	Human DNA for a no
c 23	30.6	16.7	5216	24	ABQ66852	Human polynucleoti
24	30.2	16.5	433	22	AAI89606	Human polynucleoti
25	29.8	16.3	13852	22	ABA17744	Human nervous syst
26	29.8	16.3	13852	22	ABA17745	Human nervous syst
27	29.8	16.3	13852	22	ABA21395	Human nervous syst
28	29.8	16.3	13852	22	ABA21396	Human nervous syst
29	29.6	16.2	4068	22	AAI86707	Human polynucleoti
30	29.2	16.0	2086	23	ABL12837	Drosophila melanog
31	29.2	16.0	3598	22	AAI64774	Human glucose tran
c 32	29.2	16.0	18733	22	AAK80682	Human immune/haema
33	29.2	16.0	100301	24	ABQ88176	Human osteoblast d
c 34	29	15.8	3212	24	ABK11142	cDNA of human clon
c 35	29	15.8	38886	20	AAZ23897	Murine LOBO homolo
c 36	29	15.8	49999	20	AAZ23891	Murine LOBO genomi
c 37	28.8	15.7	431	21	AAC28387	Human secreted pro
c 38	28.8	15.7	1090	18	AAT72790	Metastasis inducin
39	28.8	15.7	8894	23	ABK42719	Genomic sequence #
40	28.4	15.5	1069	24	ABK63947	cDNA encoding huma
c 41	28.4	15.5	2001	24	AAS16204	Human disulfide co
c 42	28.4	15.5	2513	24	AAS16206	Human disulfide co
43	28.4	15.5	44576	21	AAZ61522	Cosmid CVO14 conta
c 44	28.2	15.4	341	22	ABA16249	Human nervous syst
45	28.2	15.4	1907	20	AAV84508	Human secreted pro

ALIGNMENTS

RESULT 1
AAV38517
ID AAV38517 standard; DNA; 183 BP.
XX
AC AAV38517;
XX
DT 08-OCT-1998 (first entry)
XX
DE DNA marker of metastatic prostate cancer, UC Band#321.
XX
KW DNA marker; metastatic prostate cancer; human; UC Band#321; detection;
KW disease marker identification; lupus erythematosus; rheumatoid arthritis;
KW multiple sclerosis; asthma; myasthenia gravis; autoimmune thyroiditis;
KW amyloid lateral sclerosis; interstitial cystitis; prostatis; ss.
XX
OS Homo sapiens.
XX
PN WO9824935-A1.
XX
PD 11-JUN-1998.
XX
PF 05-DEC-1997; 97WO-US22105.
XX
PR 24-MAR-1997; 97US-0041576.
PR 06-DEC-1996; 96US-0032619.
PR 12-DEC-1996; 96US-0032701.
XX
PA (UROC-) UROCOR INC.
XX
PI An G, O'Hara M, Ralph D, Veltri R;
XX
DR WPI; 1998-333350/29.

```
XX
PT   Identifying markers for disease states - by amplifying RNA from
PT   peripheral blood and identifying RNA which is differential expressed
PT   between normal and disease state subjects
XX
PS   Claim 17; Page 92; 158pp; English.
XX
CC   This sequence represents a DNA marker of metastatic prostate cancer,
CC   designated UC Band#321, and was identified using a method of the
CC   invention. The method is for identifying markers for a disease state, and
CC   comprises: (a) providing a first set of peripheral blood mRNAs from one
CC   or more subjects known to exhibit the disease state and a second set of
CC   peripheral blood mRNAs from one or more normal subjects; (b) amplifying
CC   both sets of mRNAs to provide nucleic acid amplification products;
CC   (c) comparing the sets of amplification products; and (d) identifying
CC   those mRNAs that are differentially expressed between normal subjects and
CC   subjects exhibiting the disease state; where a difference in quantity of
CC   expression of an mRNA is indicative of a disease marker. This marker
CC   sequence can be used in a method of detecting a metastatic cancer disease
CC   state, especially for detection prostate cancer. Using the methods, a
CC   disease state may be detected, diagnosed, or a prognosis may be delivered
CC   by examining a blood sample rather than relying on a more invasive, or
CC   less sensitive test. In addition, a subject may be monitored for disease
CC   progression, status and response to therapies through monitoring of
CC   differentially expressed disease markers. The methods can be used for
CC   diseases such as cancer (especially metastatic or prostate cancer),
CC   asthma, lupus erythematosus, rheumatoid arthritis, multiple sclerosis,
CC   myasthenia gravis, autoimmune thyroiditis, amyloid lateral sclerosis,
CC   interstitial cystitis, prostatis or other systemic or chronic conditions.
XX
SQ   Sequence 183 BP; 43 A; 52 C; 45 G; 43 T; 0 other;

  Query Match             100.0%;  Score 183;  DB 19;  Length 183;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-53;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60

Qy      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120

Qy     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180

Qy     181 CCA 183
             |||
Db     181 CCA 183


RESULT 2
AAZ30719
ID   AAZ30719 standard; cDNA; 183 BP.
XX
AC   AAZ30719;
XX
DT   19-JAN-2000  (first entry)
XX
DE   Human UC Band #321 cDNA.
XX
KW   UC Band #321; marker; expression; diagnosis;
KW   differential; disease; cancer; metastatic; breast cancer; prostate;
KW   peripheral leukocyte; immune response; asthma; lupus erythematosus;
KW   rheumatoid arthritis; multiple sclerosis; myasthenia gravis;
KW   autoimmune thyroiditis; amyotrophic lateral sclerosis; ALS;
KW   interstitial cystitis; prostatitis; mRNA; reverse transcriptase PCR;
KW   RT-PCR; screening; early; diagnosis; prognosis; monitoring; ss.
XX
OS   Homo sapiens.
XX
PN   WO9949083-A1.
XX
PD   30-SEP-1999.
XX
PF   24-MAR-1999;   99WO-US06488.
XX
PR   24-MAR-1998;   98US-0046894.
XX
PA   (UROC-) UROCOR INC.
XX
PI   Ralph D,  An G,  O'Hara SM,  Veltri RW;
XX
DR   WPI; 1999-591105/50.
XX
PT   Identifying markers of human disease, specifically for diagnosis of
PT   metastatic prostatic and breast cancers  -
XX
PS   Claim 17; Page 124-125; 225pp; English.
XX
CC   This sequence represents a human cDNA sequence designated UC Band #321,
CC   which encodes a previously undescribed gene product. The expression of
CC   this gene in peripheral leukocyteswas examined using reverse
CC   transcriptase-PCR (RT-PCR) primers AAZ30739-Z30740. This gene was found
CC   to be differentially expressed between healthy subjects and patients
CC   with metastatic cancers (especially those of the prostate or breast) and
CC   may therefore be used as a marker for such diseases. Detecting
CC   levels of such human disease markers is used for diagnosis (also
CC   prognosis and monitoring) of diseases, including metastatic or
CC   organ-confined cancers, and diseases which also elicit an immune
CC   response such as asthma, lupus erythematosus, rheumatoid arthritis,
CC   multiple sclerosis, myasthenia gravis, autoimmune thyroiditis,
CC   amyotrophic lateral sclerosis (ALS), interstitial cystitis and
CC   prostatitis, but especially metastatic prostatic and breast cancer. A
CC   particular use is differentiating between prostatic cancer and benign
CC   prostatic hypertrophy, and between advanced and localised prostatic
CC   cancer, by multivariate analysis of several different markers. Cancers
CC   can be treated by administering sequences antisense to sequences that
CC   encode human disease markers. This method detects a leukocyte response
CC   to disease rather than products of diseased cells, so is suitable for
CC   large-scale screening of asymptomatic subjects. Disease can be detected
CC   at an early stage, when few, if any, diseased cells are present in the
CC   circulation. Analysis of blood samples eliminates the need for more
CC   invasive methods for obtaining samples.
XX
SQ   Sequence 183 BP; 43 A; 52 C; 45 G; 43 T; 0 other;

  Query Match             100.0%;  Score 183;  DB 20;  Length 183;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-53;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60

Qy      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120

Qy     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180

Qy     181 CCA 183
             |||
Db     181 CCA 183


RESULT 3
AAS77249
ID   AAS77249 standard; cDNA; 2113 BP.
XX
AC   AAS77249;
XX
```

```
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #13053.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG13062.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 13053; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2113 BP; 761 A; 474 C; 421 G; 457 T; 0 other;

  Query Match           22.4%;  Score 41;  DB 23;  Length 2113;
  Best Local Similarity   64.2%;  Pred. No. 0.00086;
  Matches   79;  Conservative    0;  Mismatches   40;  Indels    4;  Gaps    1;

Qy      30 GGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATT 89
           |||  |||   |    ||    |||||| ||| | |||||||||    | |   | |||
Db     170 GGCCAATCGACAGCACTGGAAGGAGATCAAAGGGTAGAAGGAGA----GATTCAGGTATC 225

Qy      90 TATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGA 149
           | ||||||  ||  ||||||||||||||||||| |   || ||||||   || | ||
Db     226 TGTTCCCCTGGTCACTTCCCTGCTGGGCTATGGCTTGGCACTGGCTGGGATCCTTTACCT 285

Qy     150 AAG 152
           |||
Db     286 AAG 288
```

```
RESULT 4
ABA51400
ID   ABA51400 standard; DNA; 164 BP.
XX
AC   ABA51400;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #10095.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes  -
XX
PS   Claim 4; SEQ ID NO 10095; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 164 BP; 39 A; 39 C; 44 G; 42 T; 0 other;

  Query Match           19.2%;  Score 35.2;  DB 22;  Length 164;
  Best Local Similarity   71.9%;  Pred. No. 0.031;
  Matches   46;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           ||||||  || |||||| |||||||  |||||| ||||||||||   |  |  | | ||
Db     100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy     112 CTGG 115
           ||
Db     160 TGGG 163
```

```
RESULT 5
ABA69411
ID   ABA69411 standard; DNA; 164 BP.
XX
AC   ABA69411;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #17716.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 4; SEQ ID NO 17716; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 164 BP; 39 A; 39 C; 44 G; 42 T; 0 other;

  Query Match           19.2%;  Score 35.2;  DB 22;  Length 164;
  Best Local Similarity   71.9%;  Pred. No. 0.031;
  Matches   46;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           ||||||  || ||||||  |||||||  ||||||  ||||||||||||   |  |  | | ||
Db     100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy     112 CTGG 115
             ||
Db     160 TGGG 163


RESULT 6
ABA36346
ID   ABA36346 standard; DNA; 164 BP.
XX
AC   ABA36346;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #14812 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 14812; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 164 BP; 39 A; 39 C; 44 G; 42 T; 0 other;

  Query Match           19.2%;  Score 35.2;  DB 22;  Length 164;
  Best Local Similarity   71.9%;  Pred. No. 0.031;
  Matches   46;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           ||||||  || ||||||  |||||||  ||||||  ||||||||||||   |  |  | | ||
Db     100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy     112 CTGG 115
             ||
Db     160 TGGG 163


RESULT 7
AAK17679
ID   AAK17679 standard; DNA; 164 BP.
XX
AC   AAK17679;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 17670.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
```

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 17670; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 164 BP; 39 A; 39 C; 44 G; 42 T; 0 other;

Query Match 19.2%; Score 35.2; DB 22; Length 164;
Best Local Similarity 71.9%; Pred. No. 0.031;
Matches 46; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
|||||| || |||||| ||||||| |||||| |||||||||||| | | | | ||
Db 100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy 112 CTGG 115
||
Db 160 TGGG 163

RESULT 8
AAK43494
ID AAK43494 standard; DNA; 164 BP.
XX
AC AAK43494;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 18051.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 18051; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 164 BP; 39 A; 39 C; 44 G; 42 T; 0 other;

Query Match 19.2%; Score 35.2; DB 22; Length 164;
Best Local Similarity 71.9%; Pred. No. 0.031;
Matches 46; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
|||||| || |||||| ||||||| |||||| |||||||||||| | | | | ||
Db 100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy 112 CTGG 115
||
Db 160 TGGG 163

RESULT 9
AAI24280
ID AAI24280 standard; DNA; 164 BP.
XX
AC AAI24280;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #14213 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

```
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 14213; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 164 BP; 39 A; 39 C; 44 G; 42 T; 0 other;

  Query Match           19.2%;  Score 35.2;  DB 22;  Length 164;
  Best Local Similarity   71.9%;  Pred. No. 0.031;
  Matches   46;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           ||||||  || |||||| |||||||  |||||| ||||||||||    |  |  | | ||
Db     100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy     112 CTGG 115
           ||
Db     160 TGGG 163


RESULT 10
ABL30203
ID   ABL30203 standard; DNA; 1524 BP.
XX
AC   ABL30203;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 42082.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 42082; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1524 BP; 455 A; 333 C; 353 G; 383 T; 0 other;

  Query Match           18.4%;  Score 33.6;  DB 23;  Length 1524;
  Best Local Similarity   53.9%;  Pred. No. 0.27;
  Matches   69;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy      38 CCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCC 97
           || || |||   | |||||    | ||||    |||| ||||     |  || |||| ||
Db     400 CCAAACGATTGTCCGAGATGTTCGGGCAGGTCCAGAAAGTGGGTGATCATTTAATTCACC 459

Qy      98 CAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTA 157
              |||    |  ||  ||| | |||||  ||  ||  |     ||| || ||| |   |
Db     460 ATTTGCTGGACAGGCAAGGCCAAGGATGTCCACAGGAAGTGGATATCCAGCAAAAACTCA 519

Qy     158 TGGCTTCT 165
              | || |
Db     520 GAGTTTAT 527


RESULT 11
ABL30202/c
ID   ABL30202 standard; DNA; 3819 BP.
XX
AC   ABL30202;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 42079.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 42079; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
```

```
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3819 BP; 1093 A; 806 C; 787 G; 1133 T; 0 other;

  Query Match             17.9%;  Score 32.8;  DB 23;  Length 3819;
  Best Local Similarity   55.2%;  Pred. No. 0.74;
  Matches   64;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy      38 CCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCC 97
           || || |||   | |||||    | ||||    |||| ||||     |  || |||| ||
Db    2420 CCAAACGATTGTCCGAGATGTTCGGGCAGGTCCAGAAAGTGGGTGATCATTTAATTCACC 2361

Qy      98 CAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGA 153
             |||    |  ||  ||| | |||||  ||  ||  |     ||| || ||| |
Db    2360 ATTTGCTGGACAGGCAAGGCCAAGGATGTCCACAGGAAGTGGATATCCAGCAAAAA 2305


RESULT 12
AAK73760
ID   AAK73760 standard; DNA; 3698 BP.
XX
AC   AAK73760;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28572.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
```

```
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 28572; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 3698 BP; 1067 A; 884 C; 788 G; 959 T; 0 other;

  Query Match             17.0%;  Score 31.2;  DB 22;  Length 3698;
  Best Local Similarity   55.6%;  Pred. No. 2.6;
  Matches   60;  Conservative    0;  Mismatches  48;  Indels    0;  Gaps    0;

Qy      74 ATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGG 133
           | || ||   |||| | | || || |||||| |   || |  |    ||  |||| || |
Db     590 AAGTAGTTCTGCTAATAAATCACCTAGTGCCCTGATTGGTATGAGCAGGGCGAACTGTAG 649

Qy     134 CTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCAC 181
           |  | |||||  |   |||||  |        | ||  |   |  ||||||
Db     650 CAAAGTTCATTTCTAAAAGAACCTAAATGGCCGTGTGTGGTGGCTCAC 697


RESULT 13
AAK83893
ID   AAK83893 standard; DNA; 3698 BP.
XX
AC   AAK83893;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38705.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
```

PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 38705; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 3698 BP; 1067 A; 884 C; 788 G; 959 T; 0 other;

Query Match 17.0%; Score 31.2; DB 22; Length 3698;
Best Local Similarity 55.6%; Pred. No. 2.6;
Matches 60; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 74 ATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGG 133
Db 590 AAGTAGTTCTGCTAATAAATCACCTAGTGCCCTGATTGGTATGAGCAGGGCGAACTGTAG 649

Qy 134 CTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCAC 181
Db 650 CAAAGTTCATTTCTAAAAGAACCTAAATGGCCGTGTGTGGTGGCTCAC 697

RESULT 14
AAK64703
ID AAK64703 standard; DNA; 10548 BP.
XX
AC AAK64703;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19515.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX

OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

```
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 19515; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 10548 BP; 3143 A; 2302 C; 2158 G; 2945 T; 0 other;

  Query Match             17.0%;  Score 31.2;  DB 22;  Length 10548;
  Best Local Similarity   55.6%;  Pred. No. 3.9;
  Matches   60;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy      74 ATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGG 133
           | || ||   |||| | | || || |||||| |   || |  |    ||  |||| || |
Db    7440 AAGTAGTTCTGCTAATAAATCACCTAGTGCCCTGATTGGTATGAGCAGGGCGAACTGTAG 7499

Qy     134 CTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCAC 181
           |  | |||||  |   |||||  |        | || |   |  ||||||
Db    7500 CAAAGTTCATTTCTAAAAGAACCTAAATGGCCGTGTGTGGTGGCTCAC 7547


RESULT 15
AAK69813
ID   AAK69813 standard; DNA; 10548 BP.
XX
AC   AAK69813;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24625.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
```

PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 24625; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 10548 BP; 3143 A; 2302 C; 2158 G; 2945 T; 0 other;

Query Match 17.0%; Score 31.2; DB 22; Length 10548;
Best Local Similarity 55.6%; Pred. No. 3.9;
Matches 60; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 74 ATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGG 133
Db 7440 AAGTAGTTCTGCTAATAAATCACCTAGTGCCCTGATTGGTATGAGCAGGGCGAACTGTAG 7499

Qy 134 CTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCAC 181
Db 7500 CAAAGTTCATTTCTAAAAGAACCTAAATGGCCGTGTGTGGTGGCTCAC 7547

Search completed: May 25, 2003, 03:13:59
Job time : 250 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:09:12 ; Search time 1754 Seconds
(without alignments)
1689.723 Million cell updates/sec

Title: US-09-660-568-49
Perfect score: 183
Sequence: 1 cacacactcccccattctga.....ctgtctcctggagctcacca 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	41.2	22.5	961	12	BG775699	BG775699 602650070
	2	37.6	20.5	501	17	BH309749	BH309749 CH230-181
	3	37	20.2	506	17	AQ419671	AQ419671 RPCI-11-1
c	4	36.6	20.0	789	17	BH317352	BH317352 CH230-119
c	5	35.4	19.3	702	12	BG719810	BG719810 602691217
	6	34.2	18.7	831	12	BF571585	BF571585 602078004
	7	34	18.6	283	12	BG161919	BG161919 dc43b12.y
	8	33.6	18.4	695	9	AI512580	AI512580 LD44491.5
c	9	33	18.0	438	9	AI243308	AI243308 qh41b03.x
c	10	33	18.0	467	9	AI806410	AI806410 wf27f08.x
	11	33	18.0	481	17	BH274856	BH274856 CH230-48J
	12	33	18.0	594	17	BH268816	BH268816 CH230-186
c	13	32.8	17.9	582	13	BM663839	BM663839 UI-E-CL1-
c	14	32.8	17.9	994	17	CNS0174A	AL107572 Drosophil
c	15	32.4	17.7	226	17	AQ543989	AQ543989 RPCI-11-3
c	16	32.4	17.7	524	17	B01998	B01998 cSRL-145D12
	17	32.4	17.7	544	17	AQ809155	AQ809155 HS_5323_B
c	18	32.4	17.7	609	12	BG720090	BG720090 602691009
	19	32	17.5	365	17	AZ745700	AZ745700 RPCI-24-1
	20	32	17.5	401	17	BH305232	BH305232 CH230-98C
	21	31.8	17.4	455	17	AQ536443	AQ536443 RPCI-11-3
c	22	31.8	17.4	591	12	BG606222	BG606222 WHE2962_B
	23	31.6	17.3	638	17	AG034524	AG034524 Pan trogl
c	24	31.6	17.3	649	13	BM416296	BM416296 OP21386 M
	25	31.6	17.3	720	17	BH307258	BH307258 CH230-10L
c	26	31.4	17.2	258	14	H32274	H32274 EST107204 R
	27	31.4	17.2	308	17	AQ252041	AQ252041 HS_3024_B
	28	31.4	17.2	574	13	BI359567	BI359567 384147 MA
	29	31.4	17.2	623	10	BE082216	BE082216 QV2-BT063
	30	31.4	17.2	634	9	AL510875	AL510875 AL510875
	31	31.4	17.2	695	17	BH364239	BH364239 CH230-60C
c	32	31.2	17.0	316	17	AQ584142	AQ584142 RPCI-11-4
	33	31.2	17.0	446	17	AQ977475	AQ977475 RPCI-23-3
c	34	31.2	17.0	524	17	AQ177467	AQ177467 HS_2203_B
	35	31.2	17.0	535	17	AQ284523	AQ284523 RPCI11-77
	36	31.2	17.0	751	17	AZ633721	AZ633721 1M0489C11
c	37	31	16.9	502	17	AZ217607	AZ217607 Sheared D
c	38	31	16.9	566	17	AZ792647	AZ792647 2M0045F14
c	39	31	16.9	1082	17	CNS078VV	AL434513 T3 end of
c	40	30.8	16.8	417	12	BG190341	BG190341 RST9404 A
	41	30.8	16.8	425	17	BH258367	BH258367 CH230-61A
c	42	30.8	16.8	481	17	BH860938	BH860938 UP_294-17
	43	30.8	16.8	514	17	AQ792988	AQ792988 HS_5304_A
	44	30.8	16.8	517	14	BQ090484	BQ090484 rc66b03.y
	45	30.8	16.8	577	17	AQ189764	AQ189764 HS_3224_A

ALIGNMENTS

RESULT 1
BG775699/c
LOCUS BG775699 961 bp mRNA linear EST 15-MAY-2001
DEFINITION 602650070T1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760993 3', mRNA sequence.
ACCESSION BG775699
VERSION BG775699.1 GI:14046016
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 961)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1612 row: 1 column: 18
High quality sequence start: 4
High quality sequence stop: 783.
FEATURES Location/Qualifiers

```
    source          1. .961
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4760993"
                    /clone_lib="NIH_MGC_40"
                    /tissue_type="carcinoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                    Site_2: EcoRI; cDNA made by oligo-dT priming.
                    Directionally cloned into EcoRI/XhoI sites using the
                    following 5' adaptor: GGCACGAG(G). Library constructed by
                    Ling Hong in the laboratory of Gerald M. Rubin (University
                    of California, Berkeley) using ZAP-cDNA synthesis kit
                    (Stratagene) and Superscript II RT (Life Technologies).
                    Note: this is a NIH_MGC Library."
BASE COUNT      211 a    268 c    252 g    230 t
ORIGIN

  Query Match             22.5%;  Score 41.2;  DB 12;  Length 961;
  Best Local Similarity   56.7%;  Pred. No. 0.057;
  Matches   76;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy       9 CCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAA 68
           ||| | | ||| |  |||   |||||||||| |   |   |  |||||||||| | |  |
Db     960 CCCACCTCCTGGGTTCCACCTGGCTCATCCCCACACAACGCATGAGATCCAAGAGAATGA 901

Qy      69 GGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAAC 128
           || || |  ||  |   || |       | ||  |   || ||| | |  ||   ||
Db     900 GGCGATTCGGGGCTGCTGATGTTACTGAAGACGGCACTTTTTGGTGGCCGAACGAGTGAC 841

Qy     129 AGTGGCTGACTTCA 142
           | |||| |||||||
Db     840 ATTGGCAGACTTCA 827


RESULT 2
BH309749
LOCUS       BH309749                501 bp    DNA     linear   GSS 03-DEC-2001
DEFINITION  CH230-181E7.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-181E7, DNA sequence.
ACCESSION   BH309749
VERSION     BH309749.1  GI:17235218
KEYWORDS    GSS.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-181E7.TVB
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 181  row: E  column: 7
            Seq primer: SP6
            Class: BAC ends.
FEATURES            Location/Qualifiers
     source         1. .501
                    /organism="Rattus norvegicus"
                    /strain="BN/SsNHsd/MCW"
                    /db_xref="taxon:10116"
                    /clone="CH230-181E7"
                    /clone_lib="CHORI-230 Segment 1"
                    /sex="Female"
                    /cell_type="Brain"
                    /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                    CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                    Pieter de Jong"
BASE COUNT      136 a    150 c    135 g     80 t
ORIGIN

  Query Match             20.5%;  Score 37.6;  DB 17;  Length 501;
  Best Local Similarity   63.0%;  Pred. No. 0.53;
  Matches   58;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      88 TTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAG 147
             ||| | |  | |||| | | |||| ||  || ||||   |||| ||| || ||| |||||
Db     114 TTTCTGCTGCAAGTGACCTGCCTGGTGAACTCTGGACACACAGAGGCAGAATTCCTCTAG 173

Qy     148 GAAAGAGCTATGGCTTCTGTCTCCTGGAGCTC 179
           ||  | ||  |  ||     |  || ||||| |
Db     174 GACCGGGCACTTCCTGTGTTTACCAGGAGCCC 205


RESULT 3
AQ419671
LOCUS       AQ419671                506 bp    DNA     linear   GSS 23-MAR-1999
DEFINITION  RPCI-11-179B12.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-179B12
            , DNA sequence.
ACCESSION   AQ419671
VERSION     AQ419671.1  GI:4477395
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 506)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-179B12.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES            Location/Qualifiers
     source         1. .506
                    /organism="Homo sapiens"
                    /db_xref="GDB:7568387"
                    /db_xref="taxon:9606"
                    /clone="RPCI-11-179B12"
                    /clone_lib="RPCI-11"
                    /sex="Male"
                    /cell_type="Lymphocytes"
                    /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                    RPCI11 Human Male BAC Library"
BASE COUNT      106 a    121 c    110 g    169 t
ORIGIN
```

```
  Query Match            20.2%;  Score 37;  DB 17;  Length 506;
  Best Local Similarity  71.0%;  Pred. No. 0.81;
  Matches   49;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      89 TTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGG 148
           ||  | || |||  |||||||||||||| ||||||||  ||||||||   || |  |
Db     321 TTTATTCCTCAGCTCCTTCCCTGCTGGGCCATGGATGAGTAGTGGCTGTACTCTTTAACT 380

Qy     149 AAAGAGCTA 157
           ||||  ||
Db     381 GAAGATGTA 389


RESULT 4
BH317352/c
LOCUS       BH317352                789 bp    DNA     linear   GSS 03-DEC-2001
DEFINITION  CH230-119L10.TVB CHORI-230 Segment 1 Rattus norvegicus genomic
            clone CH230-119L10, DNA sequence.
ACCESSION   BH317352
VERSION     BH317352.1  GI:17248066
KEYWORDS    GSS.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 789)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-119L10.TJB
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 119  row: L  column: 10
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .789
                     /organism="Rattus norvegicus"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-119L10"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      121 a    191 c    217 g    260 t
ORIGIN

  Query Match            20.0%;  Score 36.6;  DB 17;  Length 789;
  Best Local Similarity  60.6%;  Pred. No. 1.4;
  Matches   60;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy      85 CTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATC 144
           || ||| | |  | || | | | |||| ||  ||  |||   |||| ||| || ||| ||
Db     645 CTTTTTCTGCTGCAAGAGACCTGCCTGGTGAACTCGGGACACACAGAGGCAGAATTCCTC 586

Qy     145 TAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCACCA 183
           |||||  | ||  |  ||   ||  || ||||  |  ||
Db     585 TAGGACCGGGCATTTCCTGTGGTTACCGGGAGTCCCACA 547


RESULT 5
BG719810/c
LOCUS       BG719810                702 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602691217F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4823333 5',
            mRNA sequence.
ACCESSION   BG719810
VERSION     BG719810.1  GI:13998997
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10733  row: b  column: 06
            High quality sequence stop: 683.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4823333"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                     ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3',
                     size-selected for  average insert size 2.2 kb and
                     normalized to ROT 5. This is a primary library enriched
                     for full-length clones and constructed using the
                     Cap-trapper method (Carninci, in preparation).  Library
                     constructed by M. Brownstein (NIMH/NHGRI, National
                     Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      137 a    208 c    238 g    119 t
ORIGIN

  Query Match            19.3%;  Score 35.4;  DB 12;  Length 702;
  Best Local Similarity  53.2%;  Pred. No. 3.1;
  Matches   75;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy      25 CAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGG 84
           | | ||| |||    ||| |||   | | | |  | |   |||       ||    | |
Db     663 CTACAGGATCACTGTTAATGATTCTCCGTGCTGTAGGCTGAGAGCTTCTGTGCCACGTGC 604

Qy      85 CTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATC 144
           |  |  ||  || |  ||||      |||||   |||  |||| | ||| || || | ||
Db     603 CCTTCCATCACCGCTTTGCCAGAGAGGCTGGCACATGTCTGAAGACTGGTTGGCTCCTTC 544

Qy     145 TAGGAAAGAGCTATGGCTTCT 165
           | |||| |||  |  | ||||
Db     543 TGGGAAGGAGTGACAGGTTCT 523


RESULT 6
BF571585
LOCUS       BF571585                831 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602078004F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4252465 5',
```

```
            mRNA sequence.
ACCESSION   BF571585
VERSION     BF571585.1  GI:11645297
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1077  row: h  column: 02
            High quality sequence stop: 571.
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4252465"
                     /clone_lib="NIH_MGC_62"
                     /tissue_type="melanotic melanoma, high MDR"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.75 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      219 a    190 c    216 g    206 t
ORIGIN

  Query Match             18.7%;  Score 34.2;  DB 12;  Length 831;
  Best Local Similarity   58.3%;  Pred. No. 7.9;
  Matches   60;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy      64 CAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGA 123
           ||||  ||| |    |||  |||||| ||||| ||| || ||    | | ||  | |  ||
Db     334 CAGATCGAGGACTGTGTGGTGCTATTCATTCCTCCATTGGCTAAACAGATGAACAAGCGA 393

Qy     124 TGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTG 166
           |       || ||  || || |||||||  |   |||| ||
Db     394 GGTTGCTACACTAACAGCAGCTGGGAAAGAAGTTATGCTTGTG 436


RESULT 7
BG161919
LOCUS       BG161919                283 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  dc43b12.y1 NICHD XGC Emb3 Xenopus laevis cDNA clone IMAGE:3399886
            5', mRNA sequence.
ACCESSION   BG161919
VERSION     BG161919.1  GI:12695838
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 283)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: dc43b12.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 269.
FEATURES             Location/Qualifiers
     source          1. .283
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:3399886"
                     /clone_lib="NICHD XGC Emb3"
                     /tissue_type="embryo (stages 24-25)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 1.7 kb. Constructed by Life Technologies. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT       76 a     58 c     68 g     81 t
ORIGIN

  Query Match             18.6%;  Score 34;  DB 12;  Length 283;
  Best Local Similarity   61.1%;  Pred. No. 5.1;
  Matches   55;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy      63 GCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGG 122
           ||||| | ||| ||||    | ||||||||||  ||| | |    || || ||||
Db     157 GCAGATGAAGATTGTGTAAGGCCTATTTATTCCAGCAGAGGCAATGCTCCTTGGCTTCAT 216

Qy     123 ATGAACAGTGGCTGACTTCATCTAGGAAAG 152
           ||    || ||||   |  | | ||||| |
Db     217 ATTGCTAGAGGCTTGTTAGACCAAGGAAGG 246


RESULT 8
AI512580
LOCUS       AI512580                695 bp    mRNA    linear   EST 23-APR-2001
DEFINITION  LD44491.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
            melanogaster cDNA clone LD44491 5 similar to Cyp310a1: FBan0010391
            'cytochrome P450' located on: 2L 37A3-37A3;: 04/10/2001, mRNA
            sequence.
ACCESSION   AI512580
VERSION     AI512580.2  GI:13770009
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 695)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     On Mar 16, 1999 this sequence version replaced gi:4421998.
            Other_ESTs: LD44491.3prime
            Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003659: arm:2L [18246373,18507269]
            estimated-cyto:36F7-37A4: 04/10/2001
            Plate: LD.444  row: H  column: 7
```

```
           High quality sequence stop: 585
           POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .695
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LD44491"
                     /clone_lib="LD Drosophila melanogaster embryo pOT2"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="XL1 Blue"
                     /note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2. "
BASE COUNT      171 a    175 c    176 g    173 t
ORIGIN

  Query Match             18.4%;  Score 33.6;  DB 9;  Length 695;
  Best Local Similarity   53.9%;  Pred. No. 11;
  Matches   69;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy      38 CCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCC 97
           || || |||   | |||||    | ||||    |||| ||||     |  || |||| ||
Db     400 CCAAACGATTGTCCGAGATGTTCGGGCAGGTCCAGAAAGTGGGTGATCATTTAATTCACC 459

Qy      98 CAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTA 157
             |||    |  ||  ||| | |||||  ||  ||  |     ||| || ||| |   |
Db     460 ATTTGCTGGACAGGCAAGGCCAAGGATGTCCACAGGAAGTGGATATCCAGCAAAAACTCA 519

Qy     158 TGGCTTCT 165
           | || |
Db     520 GAGTTTAT 527


RESULT 9
AI243308/c
LOCUS       AI243308                438 bp    mRNA    linear   EST 21-DEC-1998
DEFINITION  qh41b03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1847213 3', mRNA sequence.
ACCESSION   AI243308
VERSION     AI243308.1  GI:3838705
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 623   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 415.
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1847213"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      156 a     85 c     78 g     119 t
ORIGIN

  Query Match             18.0%;  Score 33;  DB 9;  Length 438;
  Best Local Similarity   57.1%;  Pred. No. 13;
  Matches   60;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy      41 AAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAG 100
           ||||   |   |  ||   ||    | ||||||  |   || |    |||||||||  |  |
Db     202 AAGGTACTAGTGGTATAAGAGGATGGGAGGAGAGTAAAGTTAAATCATTTATTCCTCTGG 143

Qy     101 TGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCT 145
           |||| |||    |||||||| |   ||||||  |   ||  |||
Db     142 CTCCTTTCCTTTCTGGCTATGGGTTTGCAGTGGTTTTGTTTTTCT 98


RESULT 10
AI806410/c
LOCUS       AI806410                467 bp    mRNA    linear   EST 19-DEC-1999
DEFINITION  wf27f08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2356839 3', mRNA sequence.
ACCESSION   AI806410
VERSION     AI806410.1  GI:5392976
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 619   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 415.
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2356839"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      167 a     88 c     81 g     131 t
ORIGIN

  Query Match             18.0%;  Score 33;  DB 9;  Length 467;
  Best Local Similarity   57.1%;  Pred. No. 14;
  Matches   60;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy      41 AAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAG 100
           ||||   |   |  ||   ||    | ||||||  |   || |    |||||||||  |  |
Db     203 AAGGTACTAGTGGTATAAGAGGATGGGAGGAGAGTAAAGTTAAATCATTTATTCCTCTGG 144
```

```
Qy    101 TGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCT 145
          |||| |||    |||||||| |   |||||| |   ||  |||
Db    143 CTCCTTTCCTTTCTGGCTATGGGTTTGCAGTGGTTTTGTTTTTCT 99


RESULT 11
BH274856
LOCUS       BH274856                481 bp    DNA     linear   GSS 30-NOV-2001
DEFINITION  CH230-48J1.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-48J1, DNA sequence.
ACCESSION   BH274856
VERSION     BH274856.1  GI:17187258
KEYWORDS    GSS.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-48J1.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 48  row: J  column: 1
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .481
                     /organism="Rattus norvegicus"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-48J1"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      129 a    141 c    125 g     86 t
ORIGIN

  Query Match             18.0%;  Score 33;  DB 17;  Length 481;
  Best Local Similarity   60.7%;  Pred. No. 14;
  Matches   54;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     88 TTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAG 147
          ||| | |  | || | | | |||| ||  ||  |||   |||| ||| || ||| |||||
Db    265 TTTCTGCTGCAAGAGACCTGCCTGGTGAACTCAGGACACACAGAGGCAGAATTCCTCTAG 324

Qy    148 GAAAGAGCTATGGCTTCTGTCTCCTGGAG 176
          ||  | ||  |  ||    |  || ||||
Db    325 GACCGGGCACTTCCTGTGTTTACCAGGAG 353


RESULT 12
BH268816
LOCUS       BH268816                594 bp    DNA     linear   GSS 30-NOV-2001
DEFINITION  CH230-186E20.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-186E20, DNA sequence.
ACCESSION   BH268816
VERSION     BH268816.1  GI:17181122
KEYWORDS    GSS.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 594)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
            ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
            Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 186  row: E  column: 20
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .594
                     /organism="Rattus norvegicus"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-186E20"
                     /clone_lib="CHORI-230 Segment 1"
                     /sex="Female"
                     /cell_type="Brain"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      158 a    187 c    155 g     94 t
ORIGIN

  Query Match             18.0%;  Score 33;  DB 17;  Length 594;
  Best Local Similarity   60.7%;  Pred. No. 16;
  Matches   54;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     88 TTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAG 147
          ||| | |  | || | | | |||| ||  ||  |||   |||| ||| || ||| |||||
Db    235 TTTCTGCTGCAAGAGACCTGCCTGGTGAACTCAGGACACACAGAGGCAGAATTCCTCTAG 294

Qy    148 GAAAGAGCTATGGCTTCTGTCTCCTGGAG 176
          ||  | ||  |  ||    |  || ||||
Db    295 GACCGGGCACTTCCTGTGTTTACCAGGAG 323


RESULT 13
BM663839/c
LOCUS       BM663839                582 bp    mRNA    linear   EST 27-FEB-2002
DEFINITION  UI-E-CL1-afe-k-03-0-UI.s1 UI-E-CL1 Homo sapiens cDNA clone
            UI-E-CL1-afe-k-03-0-UI 3', mRNA sequence.
ACCESSION   BM663839
VERSION     BM663839.1  GI:18969506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
```

```
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 11-93, >ALU (matched compliment)
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="UI-E-CL1-afe-k-03-0-UI"
                     /clone_lib="UI-E-CL1"
                     /tissue_type="human retina"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-E-CL1 is a normalized cDNA library containing the
                     following tissue(s): retina. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CCGCG. This library was
                     created for the program, Gene Discovery in the Visual
                     System, supported by National Eye Institute (NEI).
                     TAG_LIB=UI-E-CL1
                     TAG_TISSUE=human retina
                     TAG_SEQ=CCGCG"
BASE COUNT      171 a    128 c     98 g    185 t
ORIGIN

  Query Match             17.9%;  Score 32.8;  DB 13;  Length 582;
  Best Local Similarity   54.0%;  Pred. No. 18;
  Matches   67;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy       8 TCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGA 67
           |||  |||    |||  ||   || ||||   ||||  |   || | || | | ||   |
Db     149 TCCTTCATATCAAGCGACAGCTGGTTCATTAGTAAGTCTTGGCATACATGCTAATGGGCA 90

Qy      68 AGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAA 127
             | |||   ||  | |  |    ||| | ||| |||  |||   |||||| |  ||  ||
Db      89 GGCAGAGGTTGCAGTGAGTCGAGATTGCGCCACTGCACTCCAGCCTGGGCGACTGAGCAA 30

Qy     128 CAGT 131
             | |
Db      29 GACT 26


RESULT 14
CNS0174A/c
LOCUS       CNS0174A                994 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN17G20 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL107572
VERSION     AL107572.1  GI:5627876
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 994)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .994
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN17G20"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      223 a    214 c    242 g    250 t     65 others
ORIGIN

  Query Match             17.9%;  Score 32.8;  DB 17;  Length 994;
  Best Local Similarity   55.2%;  Pred. No. 24;
  Matches   64;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy      38 CCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCC 97
           || || |||   | |||||    | ||||    |||| ||||     |  || |||| ||
Db     427 CCAAACGATTGTCCGAGATGTTCGGGCAGGTCCAGAAAGTGGGTGATCATTTAATTCACC 368

Qy      98 CAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGA 153
             |||    |  ||  ||| | |||||  ||  ||  |     ||| || ||| |
Db     367 ATTTGCTGGACAGGCAAGGCCAAGGATGTCCACAGGAAGTGGATATCCAGCAAAAA 312


RESULT 15
AQ543989/c
LOCUS       AQ543989                226 bp    DNA     linear   GSS 19-MAY-1999
DEFINITION  RPCI-11-357B13.TV RPCI-11 Homo sapiens genomic clone RPCI-11-357B13
            , DNA sequence.
ACCESSION   AQ543989
VERSION     AQ543989.1  GI:4869383
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 226)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI-11-357B13.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
```

```
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .226
                     /organism="Homo sapiens"
                     /db_xref="GDB:7636740"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-357B13"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       58 a     58 c     59 g     51 t
ORIGIN

  Query Match             17.7%;  Score 32.4;  DB 17;  Length 226;
  Best Local Similarity   64.9%;  Pred. No. 14;
  Matches   48;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy      49 CCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCC 108
           | ||||||  |||||  ||||| ||     || | |  ||| |||||||  |  || ||
Db     200 CAAGAGATGAAAGTGTGGAAGGCGAGGAAAGTCAAGGAATTCATTCCCCTGGGTCCCTCA 141

Qy     109 CTGCTGGGCTATGG 122
           |||||  ||| | |
Db     140 CTGCTTTGCTTTAG 127


Search completed: May 25, 2003, 04:15:51
Job time : 1765 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:09:47 ; Search time 52 Seconds
(without alignments)
1079.266 Million cell updates/sec

Title: US-09-660-568-49
Perfect score: 183
Sequence: 1 cacacactcccccattctga.....ctgtctcctggagctcacca 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	183	100.0	183	4	US-09-046-894-49	Sequence 49, Appl
c	2	28.2	15.4	2846	4	US-08-915-795-1	Sequence 1, Appli
c	3	28	15.3	2297	4	US-09-471-016-12	Sequence 12, Appl
c	4	27.6	15.1	2367	3	US-08-276-968A-17	Sequence 17, Appl
	5	27.4	15.0	3243	2	US-08-239-276-6	Sequence 6, Appli
	6	27.4	15.0	3243	2	US-08-468-579B-6	Sequence 6, Appli
	7	27.4	15.0	3243	3	US-08-468-577B-6	Sequence 6, Appli
c	8	27	14.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
c	9	27	14.8	246240	2	US-08-724-394A-21	Sequence 21, Appl
c	10	27	14.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
c	11	26.6	14.5	2624	1	US-08-016-863-15	Sequence 15, Appl
c	12	26.6	14.5	2624	3	US-08-276-968A-15	Sequence 15, Appl
	13	26	14.2	3138	4	US-09-434-408-1	Sequence 1, Appli
	14	26	14.2	3834	3	US-09-209-668-18	Sequence 18, Appl
	15	26	14.2	3854	1	US-08-365-470-1	Sequence 1, Appli
	16	26	14.2	3858	2	US-08-344-155C-98	Sequence 98, Appl
	17	26	14.2	3858	4	US-09-009-490A-88	Sequence 88, Appl
	18	26	14.2	3863	4	US-08-482-073-1	Sequence 1, Appli
	19	26	14.2	3863	6	5217870-1	Patent No. 5217870
	20	25.8	14.1	2380	6	5268463-1	Patent No. 5268463
c	21	25.6	14.0	1473	4	US-09-518-914-7	Sequence 7, Appli
c	22	25.4	13.9	1001	4	US-09-641-638-214	Sequence 214, App
c	23	25.4	13.9	1001	4	US-09-641-638-215	Sequence 215, App
c	24	25.4	13.9	2777	4	US-09-310-463-3	Sequence 3, Appli
c	25	25.4	13.9	2777	4	US-08-842-248A-3	Sequence 3, Appli
c	26	25.4	13.9	2790	3	US-08-985-950-21	Sequence 21, Appl
c	27	25.4	13.9	2922	4	US-09-310-463-1	Sequence 1, Appli
c	28	25.4	13.9	2922	4	US-08-842-248A-1	Sequence 1, Appli
	29	25.4	13.9	111282	4	US-09-754-250-3	Sequence 3, Appli
	30	25.2	13.8	285	4	US-09-134-001C-2573	Sequence 2573, Ap
c	31	25.2	13.8	441	4	US-09-134-001C-2432	Sequence 2432, Ap
c	32	25.2	13.8	1001	4	US-09-641-638-250	Sequence 250, App
c	33	25.2	13.8	1001	4	US-09-641-638-258	Sequence 258, App
c	34	25.2	13.8	1512	3	US-08-803-603-4	Sequence 4, Appli
c	35	25.2	13.8	1515	3	US-08-369-822C-7	Sequence 7, Appli
c	36	25.2	13.8	1515	3	US-08-582-776C-7	Sequence 7, Appli
c	37	25.2	13.8	1515	3	US-08-434-831B-7	Sequence 7, Appli
c	38	25.2	13.8	1521	3	US-08-779-764A-10	Sequence 10, Appl
c	39	25.2	13.8	1521	3	US-08-779-764A-11	Sequence 11, Appl
c	40	25.2	13.8	1521	3	US-08-779-764A-12	Sequence 12, Appl
c	41	25.2	13.8	8910	3	US-08-369-822C-19	Sequence 19, Appl
c	42	25.2	13.8	8910	3	US-08-582-776C-19	Sequence 19, Appl
c	43	25.2	13.8	8910	3	US-08-434-831B-19	Sequence 19, Appl
	44	25.2	13.8	9126	1	US-08-580-038-26	Sequence 26, Appl
	45	25.2	13.8	9126	2	US-08-639-857-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-09-046-894-49
; Sequence 49, Application US/09046894
; Patent No. 6190857
;  GENERAL INFORMATION:
;    APPLICANT:  Ralph, David
;    APPLICANT:  An, Gang
;    APPLICANT:  O'Hara, Mark S.
;    APPLICANT:  Veltri, Robert
;    TITLE OF INVENTION:  DIAGNOSIS OF DISEASE STATE USING mRNA
;    TITLE OF INVENTION:  PROFILES IN PERIPHERAL LEUKOCYTES
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/046,894
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,576
;      FILING DATE:  24-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakashima, Richard A.
;      REGISTRATION NUMBER:  P-42,023
;      REFERENCE/DOCKET NUMBER:  UROC:014
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  183 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-046-894-49

  Query Match           100.0%;  Score 183;  DB 4;  Length 183;
  Best Local Similarity  100.0%;  Pred. No. 8.6e-56;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60

Qy      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120

Qy     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180

Qy     181 CCA 183
           |||
Db     181 CCA 183
```

RESULT 2
US-08-915-795-1/c
; Sequence 1, Application US/08915795
; Patent No. 6235713
; GENERAL INFORMATION:
; APPLICANT: Marc G. ACHEN
; APPLICANT: Andrew F. WILKS
; APPLICANT: Steven A. STACKER
; APPLICANT: Kari ALITALO
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan P.L.L.C.
; STREET: 1200 G Street, NW, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,795
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2846 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Human Breast
US-08-915-795-1

Query Match 15.4%; Score 28.2; DB 4; Length 2846;
Best Local Similarity 59.3%; Pred. No. 2.7;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy       6 ACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCA 65
           ||| | ||   |||||| |  ||||  | |       ||| | ||  | |||  |||  |
Db     809 ACTGCTCCCGGCTGAGCACTGAGAGAGTAACTTTCTTGGAGGACCCCATATCGGAGTCTA 750

Qy      66 GAAGGAGAATGTGGTGAGGCT 86
```

```
           || |||| |||  ||| |||
Db     749 AAAAGAGAGTGTTCTGATGCT 729
```

RESULT 3
US-09-471-016-12/c
; Sequence 12, Application US/09471016
; Patent No. 6316195
; GENERAL INFORMATION:
; APPLICANT: Frederick, Reid D.
; APPLICANT: Tooley, Paul W.
; APPLICANT: Bonde, Morris R.
; APPLICANT: Knorr, David A.
; APPLICANT: Peterson, Gary L.
; APPLICANT: Schaad, No. 6316195man W.
; TITLE OF INVENTION: An Improved Method for the Detection and Identification
; TITLE OF INVENTION: of Tilletia indica, the Causal Organism of Karnal Bunt
; TITLE OF INVENTION: Disease by PCR
; FILE REFERENCE: 0101.98
; CURRENT APPLICATION NUMBER: US/09/471,016
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Tilletia indica
US-09-471-016-12

Query Match 15.3%; Score 28; DB 4; Length 2297;
Best Local Similarity 66.7%; Pred. No. 2.8;
Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy      19 GAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTG 78
           |||  ||  |  || | |||||| ||  | ||| || | | | |||||||| |||| | |
Db     312 GAGAGCCTGGCAGCCCTTCCCTAGGGGAGCCCACAGGTACTATTGCAGAAGAAGAAGGCG 253
```

RESULT 4
US-08-276-968A-17/c
; Sequence 17, Application US/08276968A
; Patent No. 6015701
; GENERAL INFORMATION:
; APPLICANT: Pierce, James M.
; APPLICANT: Shoreibah, Mohamed G.
; APPLICANT: Adler, Beverly
; APPLICANT: Fregien, Nevis L.
; TITLE OF INVENTION: N-Acetylglucosaminyltransferase V
; TITLE OF INVENTION: Proteins and Sequences
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,968A
; FILING DATE: 19-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/905,795
; FILING DATE: 29-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/016,863
; FILING DATE: 10-FEB-1993
; ATTORNEY/AGENT INFORMATION:

```
;       NAME:  Ferber, Donna M.
;       REGISTRATION NUMBER:  33,878
;       REFERENCE/DOCKET NUMBER:  34-92D
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303)-499-8080
;       TELEFAX:  (303)-499-8089
;       TELEX:  49617824
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2367 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  145..2367
US-08-276-968A-17

  Query Match             15.1%;  Score 27.6;  DB 3;  Length 2367;
  Best Local Similarity   48.1%;  Pred. No. 4;
  Matches   78;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy       9 CCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAA 68
           ||   ||   |  ||   || |||  ||||    |  |||   | |  ||  |     |
Db    2335 CCTTGATGAAGTCCCGGCAGGGGCAGATCCTCTGGTGTGTGGGGTGGGCCCCGGCACAAC 2276

Qy      69 GGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAAC 128
             ||  |   || ||  |  ||    | | |||||| | | ||   ||   | |||||
Db    2275 TGAAGAGCAGGAGATCCCCTTGGAGCACACAGTGCTTGCTCTTGGGGTAGAAGGATGGCA 2216

Qy     129 AGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTC 170
             || || | |  |  ||    ||||| ||||   |  |||
Db    2215 CCAGGATGTCCTTGTACAGTTCTGAGCTTTGGCAGGTCACTC 2174


RESULT 5
US-08-239-276-6
; Sequence 6, Application US/08239276
; Patent No. 5840836
;  GENERAL INFORMATION:
;    APPLICANT:  Rabin, Daniel
;    TITLE OF INVENTION:  PANCREATIC ISLET CELL ANTIGENS
;    TITLE OF INVENTION:  OBTAINED BY MOLECULAR CLONING
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sprung Kramer Schaefer & Briscoe
;      STREET:  660 White Plains Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-5144
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.5
;      SOFTWARE:  WordPerfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/239,276
;      FILING DATE:  05-MAY-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/872,646
;      FILING DATE:  08-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/715,181
;      FILING DATE:  14-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/441,703
;      FILING DATE:  04-DEC-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/312,543
;      FILING DATE:  17-FEB-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurt G. Briscoe
;      REGISTRATION NUMBER:  33,141
;      REFERENCE/DOCKET NUMBER:  MDI 251.4-KGB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (914) 332-1700
;      TELEFAX:  (914) 332-1844
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3243 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-239-276-6

  Query Match             15.0%;  Score 27.4;  DB 2;  Length 3243;
  Best Local Similarity   57.6%;  Pred. No. 5.4;
  Matches   49;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      35 ATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTC 94
             ||  | ||||    | |||||  || ||   ||||||||     | |   | |  | ||||
Db     161 ATGTCAAAGGGACCCAAGAGAGCCTAGCAGAGAAGGAGCTCCAGCTTCTGGTCATGATTC 220

Qy      95 CCCCAGTGCCTTCCCTGCTGGGCTA 119
             ||   || |  |||||| || | |
Db     221 ACCAGCTGTCCACCCTGCGGGACCA 245


RESULT 6
US-08-468-579B-6
; Sequence 6, Application US/08468579B
; Patent No. 5981700
;  GENERAL INFORMATION:
;    APPLICANT:  Rabin, Daniel
;    TITLE OF INVENTION:  PANCREATIC ISLET CELL ANTIGENS
;    TITLE OF INVENTION:  OBTAINED BY MOLECULAR CLONING
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sprung Kramer Schaefer & Briscoe
;      STREET:  660 White Plains Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-5144
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.5
;      SOFTWARE:  WordPerfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,579B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/239,276
;      FILING DATE:  05-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/872,646
;      FILING DATE:  08-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/715,181
;      FILING DATE:  14-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/441,703
;      FILING DATE:  04-DEC-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/312,543
;      FILING DATE:  17-FEB-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurt G. Briscoe
;      REGISTRATION NUMBER:  33,141
;      REFERENCE/DOCKET NUMBER:  MDI 251.5-KGB
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (914) 332-1700
;      TELEFAX:  (914) 332-1844
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3243 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-468-579B-6

  Query Match             15.0%;  Score 27.4;  DB 2;  Length 3243;
  Best Local Similarity   57.6%;  Pred. No. 5.4;
  Matches   49;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      35 ATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTC 94
           ||  | ||||    | ||||| || ||   ||||||||     | |   | |  | ||||
Db     161 ATGTCAAAGGGACCCAAGAGAGCCTAGCAGAGAAGGAGCTCCAGCTTCTGGTCATGATTC 220

Qy      95 CCCCAGTGCCTTCCCTGCTGGGCTA 119
              ||   || |  |||||| || | |
Db     221 ACCAGCTGTCCACCCTGCGGGACCA 245


RESULT 7
US-08-468-577B-6
; Sequence 6, Application US/08468577B
; Patent No. 6001804
;  GENERAL INFORMATION:
;    APPLICANT:  Rabin, Daniel
;    TITLE OF INVENTION:  PANCREATIC ISLET CELL ANTIGENS
;    TITLE OF INVENTION:  OBTAINED BY MOLECULAR CLONING
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sprung Kramer Schaefer & Briscoe
;      STREET:  660 White Plains Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-5144
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.5
;      SOFTWARE:  WordPerfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,577B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/239,276
;      FILING DATE:  05-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/872,646
;      FILING DATE:  08-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/715,181
;      FILING DATE:  14-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/441,703
;      FILING DATE:  04-DEC-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/312,543
;      FILING DATE:  17-FEB-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurt G. Briscoe
;      REGISTRATION NUMBER:  33,141
;      REFERENCE/DOCKET NUMBER:  MDI 251.8-KGB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (914) 332-1700
;      TELEFAX:  (914) 332-1844
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3243 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-468-577B-6

  Query Match             15.0%;  Score 27.4;  DB 3;  Length 3243;
  Best Local Similarity   57.6%;  Pred. No. 5.4;
  Matches   49;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      35 ATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTC 94
           ||  | ||||    | ||||| || ||   ||||||||     | |   | |  | ||||
Db     161 ATGTCAAAGGGACCCAAGAGAGCCTAGCAGAGAAGGAGCTCCAGCTTCTGGTCATGATTC 220

Qy      95 CCCCAGTGCCTTCCCTGCTGGGCTA 119
              ||   || |  |||||| || | |
Db     221 ACCAGCTGTCCACCCTGCGGGACCA 245


RESULT 8
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-20

  Query Match             14.8%;  Score 27;  DB 2;  Length 246240;
  Best Local Similarity   62.7%;  Pred. No. 54;
  Matches   42;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;
```

```
Qy      50 CAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCC 109
           ||| ||  |  ||   |||||||  | ||| ||  || ||||||||    |  |  |||
Db 111483 CAGGAATGTAGATGGCAAAGGAGAGAGAGGTTAGTGTACTTATTCCCTGCATCACCCCCC 111424

Qy     110 TGCTGGG 116
           |||| ||
Db 111423 TGCTTGG 111417


RESULT 9
US-08-724-394A-21/c
; Sequence 21, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-21

  Query Match           14.8%;  Score 27;  DB 2;  Length 246240;
  Best Local Similarity   62.7%;  Pred. No. 54;
  Matches   42;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy      50 CAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCC 109
           ||| ||  |  ||   |||||||  | ||| ||  || ||||||||    |  |  |||
Db 111483 CAGGAATGTAGATGGCAAAGGAGAGAGAGGTTAGTGTACTTATTCCCTGCATCACCCCCC 111424

Qy     110 TGCTGGG 116
           |||| ||
Db 111423 TGCTTGG 111417


RESULT 10
US-08-724-394A-22/c
; Sequence 22, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-22

  Query Match           14.8%;  Score 27;  DB 2;  Length 246240;
  Best Local Similarity   62.7%;  Pred. No. 54;
  Matches   42;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy      50 CAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCC 109
           ||| ||  |  ||   |||||||  | ||| ||  || ||||||||    |  |  |||
Db 111483 CAGGAATGTAGATGGCAAAGGAGAGAGAGGTTAGTGTACTTATTCCCTGCATCACCCCCC 111424

Qy     110 TGCTGGG 116
           |||| ||
Db 111423 TGCTTGG 111417


RESULT 11
US-08-016-863-15/c
; Sequence 15, Application US/08016863
; Patent No. 5602003
;  GENERAL INFORMATION:
;    APPLICANT:  Pierce, J. M.
;    APPLICANT:  Shoreibah, Mohamed G.
;    APPLICANT:  Adler, Beverly
;    APPLICANT:  Fregien, Nevis L.
```

```
;    TITLE OF INVENTION:  N-Acetylglucosaminyltransferase V
;    TITLE OF INVENTION:  Protein and Gene
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/016,863
;      FILING DATE:  19930210
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/905,795
;      FILING DATE:  29-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  34-92A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303)499-8080
;      TELEFAX:  (303)499-8089
;      TELEX:  823189
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  299..2521
US-08-016-863-15

  Query Match             14.5%;  Score 26.6;  DB 1;  Length 2624;
  Best Local Similarity   48.4%;  Pred. No. 9.5;
  Matches   74;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy       9 CCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAA 68
           ||   ||   |  ||   || |||  ||||    |  |||   | |  ||  |     |
Db    2489 CCTTGATGAAGTCCCGGCAGGGGCAGATCCGCTGGTGTGTGGGGTGGGCCCCGGCACAAC 2430

Qy      69 GGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAAC 128
             ||  |   || ||  |  ||    | | |||||| | | ||   ||    | ||| |
Db    2429 TGAAGAGCAGGAGATCCCCTTGGAACACACAGTGCTTGCTCTTGGGGTAGAAGGAGGGCA 2370

Qy     129 AGTGGCTGACTTCATCTAGGAAAGAGCTATGGC 161
             || || | |  | |||    ||||| ||||
Db    2369 CCAGGATGTCCTTGTATAGTTCTGAGCTTTGGC 2337


RESULT 12
US-08-276-968A-15/c
; Sequence 15, Application US/08276968A
; Patent No. 6015701
;  GENERAL INFORMATION:
;    APPLICANT:  Pierce, James M.
;    APPLICANT:  Shoreibah, Mohamed G.
;    APPLICANT:  Adler, Beverly
;    APPLICANT:  Fregien, Nevis L.
;    TITLE OF INVENTION:  N-Acetylglucosaminyltransferase V
;    TITLE OF INVENTION:  Proteins and Sequences
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/276,968A
;      FILING DATE:  19-JUL-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/905,795
;      FILING DATE:  29-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/016,863
;      FILING DATE:  10-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  34-92D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303)-499-8080
;      TELEFAX:  (303)-499-8089
;      TELEX:  49617824
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2624 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  299..2521
US-08-276-968A-15

  Query Match             14.5%;  Score 26.6;  DB 3;  Length 2624;
  Best Local Similarity   48.4%;  Pred. No. 9.5;
  Matches   74;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy       9 CCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAA 68
           ||   ||   |  ||   || |||  ||||    |  |||   | |  ||  |     |
Db    2489 CCTTGATGAAGTCCCGGCAGGGGCAGATCCGCTGGTGTGTGGGGTGGGCCCCGGCACAAC 2430

Qy      69 GGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAAC 128
             ||  |   || ||  |  ||    | | |||||| | | ||   ||    | ||| |
Db    2429 TGAAGAGCAGGAGATCCCCTTGGAACACACAGTGCTTGCTCTTGGGGTAGAAGGAGGGCA 2370

Qy     129 AGTGGCTGACTTCATCTAGGAAAGAGCTATGGC 161
             || || | |  | |||    ||||| ||||
Db    2369 CCAGGATGTCCTTGTATAGTTCTGAGCTTTGGC 2337


RESULT 13
US-09-434-408-1
; Sequence 1, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
;  APPLICANT: Venezia, Domenick
;  APPLICANT: Grossmann, Angelika
;  TITLE OF INVENTION: RING FINGER PROTEIN ZAPOP3
;  FILE REFERENCE: 98-41
;  CURRENT APPLICATION NUMBER: US/09/434,408
;  CURRENT FILING DATE: 1999-11-04
;  EARLIER APPLICATION NUMBER: US 60/108,258
```

```
;  EARLIER FILING DATE: 1998-11-12
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 3138
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (367)...(2535)
US-09-434-408-1

  Query Match           14.2%;  Score 26;  DB 4;  Length 3138;
  Best Local Similarity   52.8%;  Pred. No. 17;
  Matches   56;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy      8 TCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGA 67
          || |  | | ||||| |  |||| ||       ||  | | ||     | |||| | |
Db    473 TCTCTAAATGTGAGCTCTCAGAGATTCCATTTGGAGCTTTTGCAACATGCAAAGTTCTGC 532

Qy     68 AGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCT 113
          || |||| ||| |||   |     |.     |   |  ||||||||||
Db    533 AGAAGAAGGTGCTGATCGTCCACACGAATCACCTCACTTCCCTGCT 578


RESULT 14
US-09-209-668-18
; Sequence 18, Application US/09209668A
; Patent No. 6114517
; GENERAL INFORMATION:
;  APPLICANT: Monia, Brett P.
;  APPLICANT: Xu, Xiaoxing S.
;  TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
;  TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
;  FILE REFERENCE: ISPH-0336
;  CURRENT APPLICATION NUMBER: US/09/209,668A
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 3834
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (117)..(1949)
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: M24736/Genbank
;   DATABASE ENTRY DATE: 1994-11-07
-09-209-668-18

  Query Match           14.2%;  Score 26;  DB 3;  Length 3834;
  Best Local Similarity   49.3%;  Pred. No. 18;
  Matches   68;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy     21 GCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGT 80
          ||   ||  |   || || | |  |  || | ||||| |||| ||| |||  |    ||
Db   2072 GCTGAAACCGCAACACCCATCACCACTTCAATAGATCAAAGTCCAGCAGGCAAGGACGGC 2131

Qy     81 GAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTT 140
              |  |    | |||||    |  | ||   ||  |||| || |  |  |    |
Db   2132 CTTCAACTGAAAAGACTCAGTGTTCCCTTTCCTACTCTCAGGATCAAGAAAGTGTTGGCT 2191

Qy    141 CATCTAGGAAAGAGCTAT 158
           ||  ||| ||  | |||
Db   2192 AATGAAGGGAAAGGATAT 2209


RESULT 15
US-08-365-470-1
; Sequence 1, Application US/08365470
; Patent No. 5632991
;  GENERAL INFORMATION:
;    APPLICANT:  Gimbrone, Jr., Michael A.
;    TITLE OF INVENTION:  Antibodies Specific For E-selectin And The Uses
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX
;      STREET:  1100 New York Ave., NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/365,470
;      FILING DATE:  herewith
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/102,510
;      FILING DATE:  05-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/850,802
;      FILING DATE:  13-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Markowicz, Karen R.
;      REGISTRATION NUMBER:  36,351
;      REFERENCE/DOCKET NUMBER:  0627.1350003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3854 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-365-470-1

  Query Match           14.2%;  Score 26;  DB 1;  Length 3854;
  Best Local Similarity   49.3%;  Pred. No. 18;
  Matches   68;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy     21 GCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGT 80
          ||   ||  |   || || | |  |  || | ||||| |||| ||| |||  |    ||
Db   2072 GCTGAAACCGCAACACCCATCACCACTTCAATAGATCAAAGTCCAGCAGGCAAGGACGGC 2131

Qy     81 GAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTT 140
              |  |    | |||||    |  | ||   ||  |||| || |  |  |    |
Db   2132 CTTCAACTGAAAAGACTCAGTGTTCCCTTTCCTACTCTCAGGATCAAGAAAGTGTTGGCT 2191

Qy    141 CATCTAGGAAAGAGCTAT 158
           ||  ||| ||  | |||
Db   2192 AATGAAGGGAAAGGATAT 2209


Search completed: May 25, 2003, 04:21:50
Job time : 354 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:14:07 ; Search time 125 Seconds
(without alignments)
1933.157 Million cell updates/sec

Title: US-09-660-568-49
Perfect score: 183
Sequence: 1 cacacactcccccattctga.....ctgtctcctggagctcacca 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	35.2	19.2	164	10	US-09-864-761-21666	Sequence 21666, A
c	2	30.6	16.7	5216	9	US-10-125-540-607	Sequence 607, App
	3	30.6	16.7	5216	9	US-10-074-095-695	Sequence 695, App
c	4	30.6	16.7	5216	10	US-09-764-870-607	Sequence 607, App
	5	30.6	16.7	5216	10	US-09-764-860-695	Sequence 695, App
	6	28.8	15.7	8894	9	US-10-092-154-1606	Sequence 1606, Ap
	7	28.8	15.7	8894	10	US-09-764-847-1606	Sequence 1606, Ap
c	8	28.4	15.5	2001	9	US-09-829-155C-8	Sequence 8, Appli
c	9	28.4	15.5	2513	9	US-09-829-155C-10	Sequence 10, Appl
	10	28.2	15.4	358	9	US-10-060-036-1937	Sequence 1937, Ap
c	11	28.2	15.4	1243	12	US-10-044-090-96	Sequence 96, Appl
	12	28.2	15.4	1907	9	US-10-023-282-108	Sequence 108, App
	13	28.2	15.4	2084	9	US-10-023-282-234	Sequence 234, App
	14	28.2	15.4	2329	12	US-10-044-090-416	Sequence 416, App
	15	27.8	15.2	374	9	US-09-918-995-6295	Sequence 6295, Ap
c	16	27.8	15.2	20907	9	US-09-764-891-9966	Sequence 9966, Ap
	17	27.8	15.2	50000	9	US-10-152-724A-20	Sequence 20, Appl
c	18	27.6	15.1	312	10	US-09-783-590-11431	Sequence 11431, A
	19	27.6	15.1	945	9	US-10-023-601-43	Sequence 43, Appl
	20	27.6	15.1	1295	10	US-09-881-852-17	Sequence 17, Appl
	21	27.6	15.1	26668	10	US-09-962-832-222	Sequence 222, App
	22	27.6	15.1	39776	9	US-10-160-293-3	Sequence 3, Appli
c	23	27.4	15.0	534	10	US-09-920-300A-1518	Sequence 1518, Ap
c	24	27.4	15.0	534	12	US-10-033-528-1518	Sequence 1518, Ap
c	25	27.4	15.0	904	9	US-10-082-830-120	Sequence 120, App
c	26	27.4	15.0	1309	9	US-10-082-830-121	Sequence 121, App
c	27	27.2	14.9	8048	9	US-09-764-872-917	Sequence 917, App
c	28	27.2	14.9	16825	9	US-10-092-154-1495	Sequence 1495, Ap
c	29	27.2	14.9	16825	10	US-09-764-847-1495	Sequence 1495, Ap
	30	27	14.8	294	10	US-09-294-093B-1409	Sequence 1409, Ap
c	31	27	14.8	371	9	US-09-918-995-37556	Sequence 37556, A
	32	27	14.8	422	10	US-09-920-300A-583	Sequence 583, App
	33	27	14.8	422	12	US-10-033-528-583	Sequence 583, App
	34	27	14.8	501	10	US-09-919-580-482	Sequence 482, App
c	35	27	14.8	502	9	US-09-918-995-32131	Sequence 32131, A
c	36	27	14.8	511	9	US-09-918-995-32726	Sequence 32726, A
	37	27	14.8	66686	10	US-09-736-960-86	Sequence 86, Appl
	38	26.8	14.6	1497	9	US-10-011-582-3	Sequence 3, Appli
	39	26.8	14.6	2175	9	US-10-011-582-1	Sequence 1, Appli
	40	26.8	14.6	2223	9	US-09-891-762-3	Sequence 3, Appli
	41	26.8	14.6	2286	9	US-10-011-582-9	Sequence 9, Appli
	42	26.8	14.6	2373	9	US-10-011-582-5	Sequence 5, Appli
	43	26.8	14.6	2612	9	US-10-011-582-8	Sequence 8, Appli
	44	26.8	14.6	4586	9	US-09-891-762-1	Sequence 1, Appli
	45	26.8	14.6	5746	9	US-09-764-891-5830	Sequence 5830, Ap

ALIGNMENTS

RESULT 1
US-09-864-761-21666
; Sequence 21666, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21666
;   LENGTH: 164
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000498.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 9.00e-88
;   OTHER INFORMATION: SWISSPROT HIT: P79023, EVALUE 1.70e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF346467.1, EVALUE 1.40e-01
US-09-864-761-21666

  Query Match             19.2%;  Score 35.2;  DB 10;  Length 164;
  Best Local Similarity   71.9%;  Pred. No. 0.006;
  Matches   46;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           ||||||  || ||||||  |||||||  ||||||  ||||||||||    |  |  | | ||
Db     100 GAGATCAGAGAGCAGAAAGAGAATGAAGTGAGGATATTTATTCCCTTGGCTCTCTTCTTG 159

Qy     112 CTGG 115
             ||
Db     160 TGGG 163


RESULT 2
US-10-125-540-607/c
; Sequence 607, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ14C1
;  CURRENT APPLICATION NUMBER: US/10/125,540
;  CURRENT FILING DATE:  2002-04-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 646
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 607
;   LENGTH: 5216
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-125-540-607

  Query Match             16.7%;  Score 30.6;  DB 9;  Length 5216;
  Best Local Similarity   56.4%;  Pred. No. 0.92;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy      23 CCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGA 82
           | |  |||||| |  |   || |||    ||   |  |  ||||| | ||||  ||  |
Db    2539 CTCGGGAGGCTAAGGCAGGAGAATGGTGTGAACCCGGAAGGCAGAGGCAGAACTTGCAGT 2480

Qy      83 GGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGA 123
           |    |  ||| | ||| |||  |||  ||||||| |  ||
Db    2479 GAGCTTAGATTGCGCCACTGCACTCCAGGCTGGGCAACAGA 2439
```

```
RESULT 3
US-10-074-095-695
; Sequence 695, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008C1
;  CURRENT APPLICATION NUMBER: US/10/074,095
;  CURRENT FILING DATE:  2002-02-14
;  PRIOR APPLICATION NUMBER: 09/764,860
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/235,834
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/234,274
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/234,223
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/228,924
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: 60/224,518
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,369
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/224,519
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,964
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/241,809
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/249,299
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/236,327
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/241,785
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/244,617
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/225,268
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,368
```

; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 16.7%; Score 30.6; DB 9; Length 5216;
Best Local Similarity 56.4%; Pred. No. 0.92;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy      23 CCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGA 82
           | |  |||||| |  |    || |||     ||    |  |  ||||| | ||||  ||  |
Db    2678 CTCGGGAGGCTAAGGCAGGAGAATGGTGTGAACCCGGAAGGCAGAGGCAGAACTTGCAGT 2737

Qy      83 GGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGA 123
           |     |  ||| | ||| |||  |||  |||||||  |  ||
Db    2738 GAGCTTAGATTGCGCCACTGCACTCCAGGCTGGGCAACAGA 2778
```

RESULT 4
US-09-764-870-607/c
; Sequence 607, Application US/09764870

```
; Patent No. US20020042386A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ14
;  CURRENT APPLICATION NUMBER: US/09/764,870
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 646
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 607
;   LENGTH: 5216
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-870-607

  Query Match             16.7%;  Score 30.6;  DB 10;  Length 5216;
  Best Local Similarity   56.4%;  Pred. No. 0.92;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy      23 CCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGA 82
           | |  |||||| |  |   || |||    ||   |  |  ||||| | ||||  ||  |
Db    2539 CTCGGGAGGCTAAGGCAGGAGAATGGTGTGAACCCGGAAGGCAGAGGCAGAACTTGCAGT 2480

Qy      83 GGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGA 123
           |    |  ||| | ||| |||  |||  ||||||| |  ||
Db    2479 GAGCTTAGATTGCGCCACTGCACTCCAGGCTGGGCAACAGA 2439


RESULT 5
US-09-764-860-695
; Sequence 695, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008
;  CURRENT APPLICATION NUMBER: US/09/764,860
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1198
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
;   LENGTH: 5216
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-860-695

  Query Match             16.7%;  Score 30.6;  DB 10;  Length 5216;
  Best Local Similarity   56.4%;  Pred. No. 0.92;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy      23 CCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGA 82
           | |  |||||| |  |   || |||    ||   |  |  ||||| | ||||  ||  |
Db    2678 CTCGGGAGGCTAAGGCAGGAGAATGGTGTGAACCCGGAAGGCAGAGGCAGAACTTGCAGT 2737

Qy      83 GGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGA 123
           |    |  ||| | ||| |||  |||  ||||||| |  ||
Db    2738 GAGCTTAGATTGCGCCACTGCACTCCAGGCTGGGCAACAGA 2778


RESULT 6
US-10-092-154-1606
; Sequence 1606, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC009C1
;  CURRENT APPLICATION NUMBER: US/10/092,154
;  CURRENT FILING DATE:  2002-03-07
;  NUMBER OF SEQ ID NOS: 2003
;  Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1606
;   LENGTH: 8894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-092-154-1606

  Query Match             15.7%;  Score 28.8;  DB 9;  Length 8894;
  Best Local Similarity   58.0%;  Pred. No. 5;
  Matches   51;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      66 GAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATG 125
           || | |||   ||  | |       ||  | ||| |||  ||    |||||  |  ||
Db    3269 GAGGCAGAGGTTGCAGTGAGCTAAGATCGCACCACTGCACTCTAGCCTGGGTGACAGAGC 3328

Qy     126 AACAGTGGCTGACTTCATCTAGGAAAGA 153
           ||||  |   |||||||||||  ||||||
Db    3329 GACAGAGCAAGACTTCATCTAAGAAAGA 3356


RESULT 7
US-09-764-847-1606
; Sequence 1606, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC009
;  CURRENT APPLICATION NUMBER: US/09/764,847
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 2003
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1606
;   LENGTH: 8894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-847-1606

  Query Match             15.7%;  Score 28.8;  DB 10;  Length 8894;
  Best Local Similarity   58.0%;  Pred. No. 5;
  Matches   51;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      66 GAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGATG 125
           || | |||   ||  | |       ||  | ||| |||  ||    |||||  |  ||
Db    3269 GAGGCAGAGGTTGCAGTGAGCTAAGATCGCACCACTGCACTCTAGCCTGGGTGACAGAGC 3328

Qy     126 AACAGTGGCTGACTTCATCTAGGAAAGA 153
           ||||  |   |||||||||||  ||||||
Db    3329 GACAGAGCAAGACTTCATCTAAGAAAGA 3356


RESULT 8
US-09-829-155C-8/c
; Sequence 8, Application US/09829155C
; Patent No. US20020155561A1
; GENERAL INFORMATION:
;  APPLICANT: Thayer, Edward C.
;  TITLE OF INVENTION: Mammalian Disulfide Core Protein-4
;  FILE REFERENCE: 00-29
;  CURRENT APPLICATION NUMBER: US/09/829,155C
;  CURRENT FILING DATE:  2002-04-03
;  PRIOR APPLICATION NUMBER: 60/196,230
;  PRIOR FILING DATE: 2000-04-10
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 2001
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-829-155C-8

  Query Match             15.5%;  Score 28.4;  DB 9;  Length 2001;
```

```
  Best Local Similarity   58.1%;  Pred. No. 4.2;
  Matches   50;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      98 CAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTA 157
           ||  | |||||||||| | ||| |  | | | ||||   ||    |   ||||| || ||
Db     305 CACAGACTTCCCTGCAGAGCTGTCTACGGAGAGTGAAGGAAGGAAAGAAGGAATGATCTT 246

Qy     158 TGGCTTCTGTCTCCTGGAGCTCACCA 183
           | ||| ||  || |     ||  |||
Db     245 TAGCTCCTCACTGCCTCTACTGGCCA 220


RESULT 9
US-09-829-155C-10/c
; Sequence 10, Application US/09829155C
; Patent No. US20020155561A1
; GENERAL INFORMATION:
;  APPLICANT: Thayer, Edward C.
;  TITLE OF INVENTION: Mammalian Disulfide Core Protein-4
;  FILE REFERENCE: 00-29
;  CURRENT APPLICATION NUMBER: US/09/829,155C
;  CURRENT FILING DATE:  2002-04-03
;  PRIOR APPLICATION NUMBER: 60/196,230
;  PRIOR FILING DATE: 2000-04-10
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 2513
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-829-155C-10

  Query Match           15.5%;  Score 28.4;  DB 9;  Length 2513;
  Best Local Similarity   58.1%;  Pred. No. 4.5;
  Matches   50;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      98 CAGTGCCTTCCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTA 157
           ||  | |||||||||| | ||| |  | | | ||||   ||    |   ||||| || ||
Db    1281 CACAGACTTCCCTGCAGAGCTGTCTACGGAGAGTGAAGGAAGGAAAGAAGGAATGATCTT 1222

Qy     158 TGGCTTCTGTCTCCTGGAGCTCACCA 183
           | ||| ||  || |     ||  |||
Db    1221 TAGCTCCTCACTGCCTCTACTGGCCA 1196


RESULT 10
US-10-060-036-1937
; Sequence 1937, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF PANCREATIC CANCER
;  FILE REFERENCE: 210121.566
;  CURRENT APPLICATION NUMBER: US/10/060,036
;  CURRENT FILING DATE:  2002-01-30
;  NUMBER OF SEQ ID NOS: 4560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1937
;   LENGTH: 358
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-060-036-1937

  Query Match           15.4%;  Score 28.2;  DB 9;  Length 358;
  Best Local Similarity   61.6%;  Pred. No. 2.7;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      53 AGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGC 112
           |||    || | || |||||  || | |  || |||| | |||||| |  ||  ||||||
Db     164 AGAATGGAGAGGAGGAGGAGGCTGAGTTTGGGGTATTGAATCCCCCGGCTCCCACCCTGC 223

Qy     113 TGGGCTATGGATG 125
           |    | || ||
Db     224 AGCATCAAGGTTG 236


RESULT 11
US-10-044-090-96/c
; Sequence 96, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 96
;   LENGTH: 1243
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1404669CB1
US-10-044-090-96

  Query Match           15.4%;  Score 28.2;  DB 12;  Length 1243;
  Best Local Similarity   59.3%;  Pred. No. 4.2;
  Matches   48;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       6 ACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAGTGCA 65
           ||| | ||   |||||| |  |||| | |       ||| | ||  | |||  |||  |
Db     207 ACTGCTCCCGGCTGAGCACTGAGAGAGTAACTTTCTTGGAGGACCCCATATCGGAGTCTA 148

Qy      66 GAAGGAGAATGTGGTGAGGCT 86
           || |||| |||  ||| |||
Db     147 AAAAGAGAGTGTTCTGATGCT 127


RESULT 12
US-10-023-282-108
; Sequence 108, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/023,282
;  CURRENT FILING DATE: 2001-12-20
;  EARLIER APPLICATION NUMBER: 09/205,258
;  EARLIER FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
```

; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 108
; LENGTH: 1907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-282-108

Query Match 15.4%; Score 28.2; DB 9; Length 1907;
Best Local Similarity 61.6%; Pred. No. 4.8;
Matches 45; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      53 AGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGC 112
           |||    || | || |||||  || | |  || |||| | |||||| |  ||  ||||||
Db    1595 AGAATGGAGAGGAGGAGGAGGCTGAGTTTGGGGTATTGAATCCCCCGGCTCCCACCCTGC 1654

Qy     113 TGGGCTATGGATG 125
           |    | || ||
Db    1655 AGCATCAAGGTTG 1667
```

RESULT 13
US-10-023-282-234
; Sequence 234, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875

```
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 234
;   LENGTH: 2084
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (775)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (2080)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (2083)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-234

  Query Match             15.4%;  Score 28.2;  DB 9;  Length 2084;
  Best Local Similarity   61.6%;  Pred. No. 5;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      53 AGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGC 112
           |||    || | || |||||  || | |  || |||| | ||||||  |  ||  ||||||
Db    1780 AGAATGGAGAGGAGGAGGAGGCTGAGTTTGGGGTATTGAATCCCCCGGCTCCCACCCTGC 1839

Qy     113 TGGGCTATGGATG 125
           |    | || ||
Db    1840 AGCATCAAGGTTG 1852


RESULT 14
US-10-044-090-416
; Sequence 416, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 416
;   LENGTH: 2329
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 2398075CB1
US-10-044-090-416

  Query Match             15.4%;  Score 28.2;  DB 12;  Length 2329;
  Best Local Similarity   61.6%;  Pred. No. 5.2;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      53 AGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGC 112
           |||    || | || |||||  || | |  || |||| | ||||||  |  ||  ||||||
Db    2048 AGAATGGAGAGGAGGAGGAGGCTGAGTTTGGGGTATTGAATCCCCCGGCTCCCACCCTGC 2107

Qy     113 TGGGCTATGGATG 125
           |    | || ||
Db    2108 AGCATCAAGGTTG 2120


RESULT 15
US-09-918-995-6295
; Sequence 6295, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6295
;   LENGTH: 374
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-6295

  Query Match             15.2%;  Score 27.8;  DB 9;  Length 374;
  Best Local Similarity   51.2%;  Pred. No. 3.9;
  Matches   65;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           |||  |   |  | || | || |   |||| | |||   |  || || |   | |||||
Db      99 GAGCCCAGCGGCCGGATGCAGCAGCAGGTGGGTCTACACACACCACCCGCTGCCTCCCTC 158

Qy     112 CTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCC 171
           |  | ||  || ||| |  ||  |  |  |  ||    |   |  ||| |    ||||
Db     159 CGTGTCTGCGGCTGAGCTCTGCTTCTCGGCCCCTCACCACAGGAGTTGGACTGGTGCTCC 218

Qy     172 TGGAGCT 178
           |||||||
Db     219 AGGAGCT 225


Search completed: May 25, 2003, 04:19:26
Job time : 142 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:08:42 ; Search time 1933 Seconds
(without alignments)
2755.206 Million cell updates/sec

Title: US-09-660-568-49
Perfect score: 183
Sequence: 1 cacacactcccccattctga.....ctgtctcctggagctcacca 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		183	100.0	183	6	AR130466	AR130466 Sequence
2	c	152.2	83.2	146327	9	AL162253	AL162253 Human DNA
3	c	44.4	24.3	189509	9	AC099564	AC099564 Homo sapi
4		43.6	23.8	91639	9	AL162399	AL162399 Human DNA
5	c	43.4	23.7	167115	9	AC090093	AC090093 Homo sapi
6		43.4	23.7	173087	2	AC107318	AC107318 Sus scrof
7	c	43.4	23.7	184787	2	AC087354	AC087354 Homo sapi
8	c	43.4	23.7	185420	2	AC021973	AC021973 Homo sapi
9		42.4	23.2	205816	9	AC013751	AC013751 Homo sapi
10		41.8	22.8	174380	2	AC016075	AC016075 Homo sapi
11	c	41.8	22.8	175559	9	AC112498	AC112498 Homo sapi
12		41.8	22.8	177220	9	HS179D22	AL355433 Homo sapi
13		41.6	22.7	58693	2	AC107875	AC107875 Homo sapi
14	c	41.6	22.7	111461	9	AP005368	AP005368 Homo sapi
15	c	41.6	22.7	149126	9	AC105177	AC105177 Homo sapi
16	c	41.6	22.7	162987	2	AP005359	AP005359 Homo sapi
17	c	41.2	22.5	162126	9	AL354711	AL354711 Human DNA
18	c	41	22.4	60059	9	AL354657	AL354657 Human DNA
19	c	41	22.4	150409	9	AP003496	AP003496 Homo sapi
20	c	41	22.4	152142	9	AP003494	AP003494 Homo sapi
21		41	22.4	191896	9	AC009481	AC009481 Homo sapi
22		40.8	22.3	134339	9	AC114501	AC114501 Homo sapi
23		40.8	22.3	158513	2	AC011054	AC011054 Homo sapi
24		40.8	22.3	163395	2	AC103726	AC103726 Homo sapi
25		40.6	22.2	57000	9	AC078862	AC078862 Homo sapi
26		40.2	22.0	132470	9	HS203P18	Z97180 Human DNA s
27	c	40	21.9	150533	9	AL356215	AL356215 Human DNA
28	c	40	21.9	154067	2	AC087656	AC087656 Homo sapi
29	c	40	21.9	162851	2	AC009637	AC009637 Homo sapi
30	c	40	21.9	176425	2	AC090418	AC090418 Homo sapi
31		39.8	21.7	155147	9	HS992D9	AL008638 Human DNA
32		39.8	21.7	162249	2	AC020682	AC020682 Homo sapi
33	c	39.8	21.7	166973	9	AL451049	AL451049 Human DNA
34		39.6	21.6	107637	2	AC094679	AC094679 Rattus no
35		39.6	21.6	147301	2	AP001928	AP001928 Homo sapi
36		39.6	21.6	167376	9	AP001994	AP001994 Homo sapi
37		39.6	21.6	180402	2	AC016849	AC016849 Homo sapi
38		39.6	21.6	221091	9	AP003390	AP003390 Homo sapi
39		39	21.3	64491	9	AL365188	AL365188 Human DNA
40		39	21.3	172900	9	AC107027	AC107027 Homo sapi
41		38.4	21.0	51626	2	AC068840	AC068840 Homo sapi
42		38.4	21.0	103120	2	AC103223	AC103223 Rattus no
43		38.4	21.0	112622	2	AC010866	AC010866 Homo sapi
44	c	38.4	21.0	114800	9	AC123567	AC123567 Homo sapi
45		38.4	21.0	119234	2	AC105231	AC105231 Homo sapi

ALIGNMENTS

RESULT 1
AR130466
LOCUS AR130466 183 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 49 from patent US 6190857.
ACCESSION AR130466
VERSION AR130466.1 GI:14118791
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 183)
AUTHORS Ralph,D., An,G., O'Hara,S.Mark. and Veltri,R.
TITLE Diagnosis of disease state using MRNA profiles in peripheral leukocytes
JOURNAL Patent: US 6190857-A 49 20-FEB-2001;

```
FEATURES             Location/Qualifiers
     source          1. .183
                     /organism="unknown"
BASE COUNT       43 a     52 c     45 g     43 t
ORIGIN

  Query Match             100.0%;  Score 183;  DB 6;  Length 183;
  Best Local Similarity   100.0%;  Pred. No. 3e-49;
  Matches  183;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAA 60

Qy      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTAT 120

Qy     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGGAGCTCA 180

Qy     181 CCA 183
           |||
Db     181 CCA 183


RESULT 2
AL162253/c
LOCUS       AL162253              146327 bp    DNA     linear   PRI 21-MAR-2001
DEFINITION  Human DNA sequence from clone RP11-574F11 on chromosome 9, complete
            sequence.
ACCESSION   AL162253
VERSION     AL162253.17  GI:13677203
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 146327)
  AUTHORS   Clark,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 19, 2001 this sequence version replaced gi:12539553.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-574F11 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-574F11 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true right end of clone RP11-574F11 is at 146327 in this
            sequence. The true left end of clone RP11-635N21 is at 62948 in
            this sequence. The true right end of clone RP11-12D24 is at 100 in
            this sequence.
FEATURES             Location/Qualifiers
     source          1. .146327
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-574F11"
                     /clone_lib="RPCI-11.2"
     misc_feature    84304. .84605
                     /note="Sequence from overlapping clone bA635N21
                     (AL354744). Assembly confirmed by restriction digest
                     data."
     misc_feature    84304. .84306
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
BASE COUNT    44360 a  29767 c  29483 g  42717 t
ORIGIN

  Query Match             83.2%;  Score 152.2;  DB 9;  Length 146327;
  Best Local Similarity   92.5%;  Pred. No. 4.1e-39;
  Matches  160;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       2 ACACACTCCCCCATTCTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAG 61
           || | |  | |      ||||||||||||||||||||||||||||||||||||||||||||
Db 145681 ACTCTCCACTCTGCCTTGAGCCCCAAGAGGCTCATCCCTAAGGATGTCCAGAGATCCAAG 145622

Qy      62 TGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATG 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 145621 TGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATG 145562

Qy     122 GATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCCTGG 174
           |||||||||||||||||||||||||||||||||||||||||||||||||||| | | |
Db 145561 GATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCCCGTAG 145509


RESULT 3
AC099564/c
LOCUS       AC099564              189509 bp    DNA     linear   PRI 17-APR-2002
DEFINITION  Homo sapiens chromosome 1 clone RP11-359A17, complete sequence.
ACCESSION   AC099564 AL391540
VERSION     AC099564.2  GI:20163088
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 189509)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 189509)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2001) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 189509)
  AUTHORS   Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
            Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-2002) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Apr 17, 2002 this sequence version replaced gi:16946001.
            -------------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
```

Drafting Center: SC
---------- Project Information
Center project name: chr-1
Center clone name: RP11-359A17 (sc0138)
---------- Summary Statistics
Sequencing vector: plasmid; L08752; 91% of reads
Sequencing vector: plasmid; 9% of reads
Chemistry: Dye-terminator ET; 40% of reads
Chemistry: Dye-terminator Big Dye; 60% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 189243 bases at least Q40
Consensus quality: 189424 bases at least Q30
Consensus quality: 189508 bases at least Q20
Insert size: 189509; sum-of-contigs
Quality coverage: 6.9x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': Mapping in progress
3': RP11-455L10 AL357499, 57084-bp overlap

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI		BglII		HindIII	
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
8696	8707	2944	2929	529	<800
6	<800	2067	2050	6382	6380
6553	6369	6296	6357	512	<800
3847	3850	1756	1717	449	<800
1347	1329	1273	1295	3402	3320
1619	1607	199	<800	2513	2580
1509	1483	1163	1138	7289	7194
734	<800	5058	5198	1890	1946
96	<800	5880	5768	1165	1134
4125	4139	9625	10059	851	824
6418	6369	4576	4663	8925	8862
155	<800	5240	5198	5167	5132
3110	3153	519	<800	4215	4171
335	<800	42	<800	998	1068
12736	12524	5005	4897	869	824
429	<800	335	<800	674	<800
724	<800	1083	1078	432	<800
1210	1212	1903	1876	1765	1759
688	<800	7581	7291	2997	2956
2858	2845	277	<800	779	<800
6500	6369	1941	1876	3980	3919
115	<800	4444	4415	6832	7194
851	808	889	881	1239	1233
1104	1103	5285	5437	12095	12002
15	<800	972	988	1532	1538
202	<800	11	<800	2586	2580
10384	10323	4790	4897	6439	6380
660	<800	1180	1138	1250	1233
2553	2617	27590	27791	3830	3919
12237	11780	6201	6357	612	<800
712	<800	2762	2751	2547	2580
2156	2175	351	<800	6491	6380
559	<800	3383	3355	553	<800
12040	11780	13408	12546	475	<800
1752	1756	5980	6048	1993	1946
6078	6013	1006	1078	3063	3123
1736	1756	9250	9321	5043	4966
3432	3387	6101	6048	4148	4171
11502	11780	731	<800	269	<800
1277	1298	810	801	6471	6380
6984	6884	1065	1078	9762	9651
9648	10323	69	<800	1507	1538
2076	2175	3853	3789	302	<800
10638	11055	4709	4663	3328	3320
1147	1137	6349	6357	7670	7599
639	<800	1593	1555	2969	2956

7481	7433	5528	5768	7375	7194
587	<800	1106	1078	14428	14745
1195	1212	408	<800	9736	9651
820	808	9077	8845	4035	3919
369	<800	1402	1408	856	824
2622	2617	929	988	246	<800
5340	5311	868	881	902	824
4918	4881	1345	1295	3251	3320
444	<800			304	<800
2319	2391			858	824
2342	2391			2922	2956

Query Match 24.3%; Score 44.4; DB 9; Length 189509;
Best Local Similarity 67.0%; Pred. No. 0.0013;
Matches 63; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy      53 AGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGC 112
           | ||   |  |  | | ||||  | ||| ||  | | ||||||||  |  ||||||||||
Db   66065 AAATAAGATGGTGGCAAGAGAGGGAGGTCAGCATGTATATTCCCCAGGCTCCTTCCCTGC 66006

Qy     113 TGGGCTATGGATGAACAGTGGCTGACTTCATCTA 146
           | |||| ||| |   |||||||| | ||| ||||
Db   66005 TAGGCTCTGGGTTGGCAGTGGCTAAGTTCCTCTA 65972
```

RESULT 4
AL162399
LOCUS AL162399 91639 bp DNA linear PRI 25-OCT-2001
DEFINITION Human DNA sequence from clone RP11-545I10 on chromosome 1, complete sequence.
ACCESSION AL162399
VERSION AL162399.12 GI:16501106
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 91639)
AUTHORS Martin,S.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Oct 26, 2001 this sequence version replaced gi:15130717.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1
RP11-545I10 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-545I10 It may be shorter because we sequence overlapping sections only once, except for a short overlap.
The true left end of clone RP11-545I10 is at 1 in this sequence.
The true left end of clone RP4-681G13 is at 89640 in this sequence.
The true right end of clone RP11-576I22 is at 40989 in this sequence.
FEATURES Location/Qualifiers
source 1. .91639
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-545I10"
/clone_lib="RPCI-11.2"
misc_feature 502. .905
/note="Sequence from overlapping clone RP11-576I22 (AL354732). Assembly confirmed by restriction digest."
BASE COUNT 29501 a 16967 c 17634 g 27537 t
ORIGIN

Query Match 23.8%; Score 43.6; DB 9; Length 91639;
Best Local Similarity 67.8%; Pred. No. 0.0025;
Matches 61; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

```
Qy      63 GCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGG 122
           | || | |||||   ||   || ||||||||||| ||||  || ||||||   ||| |  |
Db    9496 GTAGGAAGAGAAAAAGGATGGGGTATTTATTCCTCCAGCTCCCTCCCTGGCAGGCCACAG 9555

Qy     123 ATGAACAGTGGCTGACTTCATCTAGGAAAG 152
           | ||||||||||||||  ||||  || |
Db    9556 GATAGCAGTGGCTGACTTTCTCTACAAAGG 9585
```

RESULT 5
AC090093/c
LOCUS AC090093 167115 bp DNA linear PRI 01-JUN-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-656A10, complete sequence.
ACCESSION AC090093
VERSION AC090093.7 GI:21281556
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 167115)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-656A10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167115)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,

Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 167115)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 167115)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (31-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 5 (bases 1 to 167115)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On May 31, 2002 this sequence version replaced gi:20389333.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L12753
Center clone name: 656_A_10

FEATURES	Location/Qualifiers
source	1. .167115
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="8"
	/map="8"
	/clone="RP11-656A10"
	/clone_lib="RPCI-11 Human Male BAC"
unsure	374. .393
	/note="single clone coverage"
unsure	471. .475
	/note="<30 qual SNGL region"
unsure	508. .512
	/note="<30 qual SNGL region"
unsure	572. .576
	/note="<30 qual SNGL region"
repeat_region	complement(642. .1043)
	/rpt_family="MER66A"
unsure	985. .990
	/note="<30 qual SNGL region"
unsure	1042. .1046
	/note="<30 qual SNGL region"
unsure	1062. .1067
	/note="<30 qual SNGL region"
unsure	1075. .1079
	/note="<30 qual SNGL region"
unsure	1102. .1110
	/note="<30 qual SNGL region"
unsure	1121. .1130
	/note="<30 qual SNGL region"
repeat_region	complement(1222. .1485)
	/rpt_family="ERVL"
repeat_region	complement(1551. .1852)
	/rpt_family="ERVL"
repeat_region	complement(2187. .2378)
	/rpt_family="L3"
repeat_region	2530. .2923
	/rpt_family="L2"
repeat_region	4193. .4411
	/rpt_family="AluSx"
repeat_region	4844. .4961
	/rpt_family="MIR"
repeat_region	complement(4950. .5023)
	/rpt_family="MIR3"
repeat_region	complement(5475. .5561)
	/rpt_family="L1PB3"

```
     repeat_region   7272. .7413
                     /rpt_family="MIR"
     repeat_region   7457. .7612
                     /rpt_family="L1ME2"
     repeat_region   7613. .7924
                     /rpt_family="AluSx"
     unsure          complement(7886. .7894)
                     /note="<30 qual SNGL region"
     repeat_region   7925. .9228
                     /rpt_family="L1ME2"
     repeat_region   complement(9283. .9460)
                     /rpt_family="MIR3"
     repeat_region   9633. .9774
                     /rpt_family="MIR3"
     repeat_region   complement(10149. .10394)
                     /rpt_family="MIR"
     repeat_region   complement(11430. .11625)
                     /rpt_family="L2"
     repeat_region   complement(12608. .12682)
                     /rpt_family="MIR3"
     repeat_region   12870. .13114
                     /rpt_family="AluSg"
     repeat_region   13115. .13148
                     /rpt_family="(TCTCTG)n"
     repeat_region   13149. .13292

  Query Match             23.7%;  Score 43.4;  DB 9;  Length 167115;
  Best Local Similarity   72.7%;  Pred. No. 0.0029;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      65 AGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGAT 124
           ||||| |||| | | | |||  |||| || ||  ||| ||||||||| |||  ||| || |
Db   40703 AGAAGCAGAAGGAGATCAGGGCATTTGTTACCAAAGTTCCTTCCCTTCTGCTCTAAGGTT 40644

Qy     125 GAACAGTGGCTGACTTC 141
           |||||||||  ||||
Db   40643 TCACAGTGGCTTTCTTC 40627


RESULT 6
AC107318
LOCUS       AC107318              173087 bp    DNA     linear   HTG 18-JAN-2002
DEFINITION  Sus scrofa clone RP44-137G11, WORKING DRAFT SEQUENCE, 4 unordered
            pieces.
ACCESSION   AC107318
VERSION     AC107318.1  GI:18201780
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Sus scrofa.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 173087)
  AUTHORS   Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q.,
            Legaspi,R., Maduro,Q.L., Maduro,V.B., Masiello,C., Maskeri,B.,
            Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A.,
            Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 173087)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     -------------- Genome Center
                Center: NIH Intramural Sequencing Center
                Center code: NISC
                Web site: http://www.nisc.nih.gov
                Contact: nisc_mouse@nhgri.nih.gov
            -------------- Project Information
                Center project name: caz
                Center clone name: 137G11
            -------------- Summary Statistics
                Sequencing vector: plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 171731 bases at least Q40
                Consensus quality: 172220 bases at least Q30
                Consensus quality: 172357 bases at least Q20
                Insert size: 143000; agarose-fp
                Insert size: 172787; sum-of-contigs
                Quality coverage: 11.24x in Q20 bases; agarose-fp
                Quality coverage: 9.30x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    16132: contig of 16132 bp in length
            *   16133    16232: gap of unknown length
            *   16233    33937: contig of 17705 bp in length
            *   33938    34037: gap of unknown length
            *   34038    98779: contig of 64742 bp in length
            *   98780    98879: gap of unknown length
            *   98880   173087: contig of 74208 bp in length.
FEATURES            Location/Qualifiers
     source         1. .173087
                    /organism="Sus scrofa"
                    /db_xref="taxon:9823"
                    /clone="RP44-137G11"
                    /clone_lib="RP44"
     misc_feature   1. .16132
                    /note="assembly_fragment"
     misc_feature   16233. .33937
                    /note="assembly_fragment"
     misc_feature   34038. .98779
                    /note="assembly_fragment
                    clone_end:SP6
                    vector_side:left"
     misc_feature   98880. .173087
                    /note="assembly_fragment
                    clone_end:T7
                    vector_side:left"
BASE COUNT    52234 a  32792 c  33848 g  53909 t    304 others
ORIGIN

  Query Match             23.7%;  Score 43.4;  DB 2;  Length 173087;
  Best Local Similarity   65.3%;  Pred. No. 0.0029;
  Matches   79;  Conservative    0;  Mismatches   41;  Indels    1;  Gaps    1;

Qy      47 GTCCAGAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTT 106
           | |  ||| ||  |||||  |||||  || ||| | | |||||||||| |   ||  | |
Db   15572 GGCAGGAGGTCAGAGTGCAAGAGGAGGGTGAGGTTAAGGTATTTATTCTCTTGGTTTC-T 15630

Qy     107 CCCTGCTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTG 166
           ||||||  |||| ||| | | ||||||  ||  ||| | ||      |  | |||||||
Db   15631 CCCTGCCAGGCTGTGGGTTAGCAGTGGTTGTGTTCCTTTACTCTCAATTTGTGGCTTCTC 15690

Qy     167 T 167
           |
Db   15691 T 15691


RESULT 7
AC087354/c
LOCUS       AC087354              184787 bp    DNA     linear   HTG 23-AUG-2002
DEFINITION  Homo sapiens chromosome 8 clone RP11-398H6 map 8, *** SEQUENCING IN
            PROGRESS ***, 2 ordered pieces.
```

```
ACCESSION   AC087354
VERSION     AC087354.9  GI:22450687
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 184787)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-398H6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184787)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 184787)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Aug 23, 2002 this sequence version replaced gi:22297025.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L11709
                Center clone name: 398_H_6
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1    79726: contig of 79726 bp in length
            *    79727 79826: gap of      100 bp
            *    79827   184787: contig of 104961 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184787
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-398H6"
                     /clone_lib="RPCI-11 Human Male BAC"
BASE COUNT    54973 a  39178 c  39320 g  51130 t    186 others
ORIGIN

  Query Match             23.7%;  Score 43.4;  DB 2;  Length 184787;
  Best Local Similarity   72.7%;  Pred. No. 0.0029;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      65 AGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGAT 124
           ||||| |||| | | | |||  |||| || ||  ||| |||||||| |||  ||| || |
Db 106374 AGAAGCAGAAGGAGATCAGGGCATTTGTTACCAAAGTTCCTTCCCTTCTGCTCTAAGGTT 106315

Qy     125 GAACAGTGGCTGACTTC 141
           |||||||||  ||||
Db 106314 TCACAGTGGCTTTCTTC 106298


RESULT 8
AC021973/c
LOCUS       AC021973              185420 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 8 clone RP11-398H6, WORKING DRAFT SEQUENCE,
            7 unordered pieces.
ACCESSION   AC021973
VERSION     AC021973.3  GI:8954353
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 185420)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185420)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 7, 2000 this sequence version replaced gi:7191944.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------- Project Information ----------
            Center project name: H_NH0398H06
            ---------- Summary Statistics  ----------
            Sequencing vector: M13; 89%
            Sequencing vector: plasmid; 11%
            Chemistry: Dye-primer ET; 89% of reads
            Chemistry: Dye-terminator Big Dye; 11% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 181383 bases at least Q40
            Consensus quality: 182825 bases at least Q30
            Consensus quality: 183463 bases at least Q20
```

```
          Insert size: 185000; agarose-fp
          Insert size: 184820; sum-of-contigs
          Quality coverage: 6.44 in Q20 bases; agarose-fp
          Quality coverage: 5.98 in Q20 bases; sum-of-contigs
          ----------------------------------------------------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 7 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1     2842: contig of 2842 bp in length
          *     2843     2942: gap of unknown length
          *     2943     6917: contig of 3975 bp in length
          *     6918     7017: gap of unknown length
          *     7018    15680: contig of 8663 bp in length
          *    15681    15780: gap of unknown length
          *    15781    30574: contig of 14794 bp in length
          *    30575    30674: gap of unknown length
          *    30675    61322: contig of 30648 bp in length
          *    61323    61422: gap of unknown length
          *    61423    93101: contig of 31679 bp in length
          *    93102    93201: gap of unknown length
          *    93202   185420: contig of 92219 bp in length.
FEATURES             Location/Qualifiers
     source          1. .185420
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /clone="RP11-398H6"
     misc_feature    1. .2842
                     /note="assembly_name:Contig8
                     clone_end:T7
                     vector_side:right"
     misc_feature    2943. .6917
                     /note="assembly_name:Contig9"
     misc_feature    7018. .15680
                     /note="assembly_name:Contig10"
     misc_feature    15781. .30574
                     /note="assembly_name:Contig11"
     misc_feature    30675. .61322
                     /note="assembly_name:Contig12"
     misc_feature    61423. .93101
                     /note="assembly_name:Contig13
                     clone_end:SP6
                     vector_side:left"
     misc_feature    93202. .185420
                     /note="assembly_name:Contig14"
BASE COUNT    53244 a  39520 c  39207 g  52848 t    601 others
ORIGIN

  Query Match             23.7%;  Score 43.4;  DB 2;  Length 185420;
  Best Local Similarity   72.7%;  Pred. No. 0.0029;
  Matches   56;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      65 AGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTATGGAT 124
           ||||| |||| | | | ||| |||| || ||  ||| |||||||| |||  ||| || |
Db 167596 AGAAGCAGAAGGAGATCAGGGCATTTGTTACCAAAGTTCCTTCCCTTCTGCTCTAAGGTT 167537

Qy     125 GAACAGTGGCTGACTTC 141
             |||||||||  ||||
Db 167536 TCACAGTGGCTTTCTTC 167520


RESULT 9
AC013751
LOCUS       AC013751               205816 bp    DNA     linear   PRI 07-MAY-2001
DEFINITION  Homo sapiens chromosome , clone RP11-298P6, complete sequence.
ACCESSION   AC013751
VERSION     AC013751.6  GI:13958504
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 205816)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome, clone RP11-298P6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 205816)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 205816)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 205816)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 6, 2001 this sequence version replaced gi:11094844.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2369
                Center clone name: 298_P_6
            --------------
FEATURES             Location/Qualifiers
     source          1. .205816
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-298P6"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   complement(891. .1189)
                     /rpt_family="AluJb"
     repeat_region   complement(1411. .1711)
                     /rpt_family="AluSq"
     repeat_region   3736. .3793
                     /rpt_family="(TCTA)n"
     repeat_region   complement(3794. .3914)
                     /rpt_family="FLAM_C"
     repeat_region   3985. .4277
                     /rpt_family="AluJb"
     repeat_region   complement(4315. .4526)
                     /rpt_family="L1MC5"
     repeat_region   4569. .4612
                     /rpt_family="CT-rich"
     repeat_region   4694. .4797
                     /rpt_family="FAM"
     repeat_region   complement(4827. .5088)
                     /rpt_family="L1MC5"
     repeat_region   5089. .5389
                     /rpt_family="AluSx"
     repeat_region   complement(5390. .6045)
                     /rpt_family="L1MC5"
     repeat_region   6107. .6364
                     /rpt_family="AluSx"
     repeat_region   6912. .6982
                     /rpt_family="(TATG)n"
     repeat_region   complement(6998. .7072)
                     /rpt_family="L1MC5"
     repeat_region   8649. .8703
                     /rpt_family="(TC)n"
     repeat_region   8703. .8734
                     /rpt_family="(CA)n"
     repeat_region   8810. .9110
                     /rpt_family="AluSx"
     repeat_region   9241. .9274
                     /rpt_family="(TTTTG)n"
     repeat_region   complement(9320. .10348)
                     /rpt_family="L1MC4a"
     repeat_region   complement(10376. .10678)
                     /rpt_family="AluSx"
     repeat_region   11057. .11087
                     /rpt_family="AT_rich"
     repeat_region   11626. .11660
                     /rpt_family="AT_rich"
     repeat_region   11850. .11956
                     /rpt_family="MIR3"
     repeat_region   12137. .12438
                     /rpt_family="AluJb"
     repeat_region   13183. .13359
                     /rpt_family="(TATATG)n"
     repeat_region   13549. .13576
                     /rpt_family="AT_rich"
     repeat_region   15401. .15450
                     /rpt_family="AT_rich"
     repeat_region   complement(15752. .15855)
                     /rpt_family="MER70B"
     repeat_region   16222. .16243
                     /rpt_family="AT_rich"
     repeat_region   complement(17418. .17722)
                     /rpt_family="AluSg"
     repeat_region   19520. .19602
                     /rpt_family="AT_rich"
     repeat_region   complement(20302. .20653)
                     /rpt_family="MLT1J1"
     repeat_region   complement(20968. .21083)
                     /rpt_family="AluJb"
     repeat_region   complement(21308. .21447)
                     /rpt_family="MER46C"
     repeat_region   21730. .21750
                     /rpt_family="AT_rich"
     repeat_region   22254. .22303
                     /rpt_family="AT_rich"
     repeat_region   complement(22554. .22678)
                     /rpt_family="FLAM_C"
     repeat_region   complement(23551. .23878)
                     /rpt_family="AluSq"
     repeat_region   24098. .24386
                     /rpt_family="MER7A"
     repeat_region   complement(24639. .24936)
                     /rpt_family="MLT1I"
     repeat_region   complement(24939. .25019)
                     /rpt_family="MER46C"
     repeat_region   26081. .26243
                     /rpt_family="MIR"
     repeat_region   complement(27395. .27861)
                     /rpt_family="L1MD1"
     repeat_region   complement(27862. .28152)
                     /rpt_family="AluSx"
     repeat_region   complement(28153. .28311)
                     /rpt_family="L1MD1"
     repeat_region   28316. .28456
                     /rpt_family="L1MD2"
     repeat_region   28457. .28491
                     /rpt_family="(CA)n"
     repeat_region   28492. .28687
                     /rpt_family="L1MD2"
     repeat_region   31292. .31336
                     /rpt_family="AT_rich"
     repeat_region   complement(32423. .32931)
                     /rpt_family="L2"

  Query Match             23.2%;  Score 42.4;  DB 9;  Length 205816;
  Best Local Similarity   63.3%;  Pred. No. 0.0061;
  Matches   81;  Conservative    0;  Mismatches   46;  Indels    1;  Gaps    1;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           || |||  |  |||| | |||| || | | ||   ||||||  |||| ||   || ||||
Db   50398 GAAATCAGAAAGCAGGAAGAGAGTGGGATCAGAGCATTTATCTCCCCTGTTATTTTCCTG 50457

Qy     112 CTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCTTCTGTCTCC 171
           |   ||  | | | | |||||||||   |  ||||  ||| | |||  ||| ||| | |
Db   50458 CCAAGCCTTAGGTTAGCAGTGGCTGCAGTTCTCTACCAAA-ATCTACAGCTCCTGACACA 50516

Qy     172 TGGAGCTC 179
           ||||  ||
Db   50517 TGGATGTC 50524


RESULT 10
AC016075
LOCUS       AC016075              174380 bp    DNA     linear   HTG 17-MAR-2000
DEFINITION  Homo sapiens clone RP11-699A7, WORKING DRAFT SEQUENCE, 12 unordered
            pieces.
ACCESSION   AC016075
```

```
VERSION     AC016075.4  GI:7259724
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 174380)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-699A7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174380)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 17, 2000 this sequence version replaced gi:6970517.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L4345
                Center clone name: 699_A_7
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 166563 bases at least Q40
                Consensus quality: 169997 bases at least Q30
                Consensus quality: 171469 bases at least Q20
                Insert size: 181000; agarose-fp
                Insert size: 173280; sum-of-contigs
                Quality coverage: 5.2 in Q20 bases; agarose-fp
                Quality coverage: 5.4 in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      377: contig of 377 bp in length
            *      378 477: gap of      100 bp
            *      478     1740: contig of 1263 bp in length
            *     1741 1840: gap of      100 bp
            *     1841     4031: contig of 2191 bp in length
            *     4032 4131: gap of      100 bp
            *     4132     9969: contig of 5838 bp in length
            *     9970 10069: gap of      100 bp
            *    10070    17346: contig of 7277 bp in length
            *    17347 17446: gap of      100 bp
            *    17447    29159: contig of 11713 bp in length
            *    29160 29259: gap of      100 bp
            *    29260    44992: contig of 15733 bp in length
            *    44993 45092: gap of      100 bp
            *    45093    63565: contig of 18473 bp in length
            *    63566 63665: gap of      100 bp
            *    63666    82201: contig of 18536 bp in length
            *    82202 82301: gap of      100 bp
            *    82302   105742: contig of 23441 bp in length
            *   105743 105842: gap of      100 bp
            *   105843   139307: contig of 33465 bp in length
            *   139308 139407: gap of      100 bp
            *   139408   174380: contig of 34973 bp in length.
FEATURES             Location/Qualifiers
     source          1. .174380
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-699A7"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .377
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    478. .1740
                     /note="assembly_fragment"
     misc_feature    1841. .4031
                     /note="assembly_fragment"
     misc_feature    4132. .9969
                     /note="assembly_fragment"
     misc_feature    10070. .17346
                     /note="assembly_fragment"
     misc_feature    17447. .29159
                     /note="assembly_fragment"
     misc_feature    29260. .44992
                     /note="assembly_fragment"
     misc_feature    45093. .63565
                     /note="assembly_fragment"
     misc_feature    63666. .82201
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    82302. .105742
                     /note="assembly_fragment"
     misc_feature    105843. .139307
                     /note="assembly_fragment"
     misc_feature    139408. .174380
                     /note="assembly_fragment"
BASE COUNT    53438 a  32639 c  33542 g  53658 t   1103 others
ORIGIN

  Query Match             22.8%;  Score 41.8;  DB 2;  Length 174380;
  Best Local Similarity   63.4%;  Pred. No. 0.0096;
  Matches   64;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           |||||  | |  |||||| | |    | | | | ||||| |||||  ||   | ||| ||
Db 154021 GAGATTGAGGGACAGAAGAATAGCAAGATCATGATATTTGTTCCCTTAGCTACATCCATG 154080

Qy     112 CTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAG 152
           | |||| |  | |   ||||| |||  ||| ||||| || |
Db 154081 CAGGGCCACTGGTTGGCAGTGTCTGCTTTCCTCTAGTAAGG 154121


RESULT 11
AC112498/c
LOCUS       AC112498              175559 bp    DNA     linear   PRI 28-JUN-2002
DEFINITION  Homo sapiens X BAC RP11-699A7 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC112498
VERSION     AC112498.3  GI:21591809
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175559)
```

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175559)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 175559)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 175559)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 175559)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jun 26, 2002 this sequence version replaced gi:20335921.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES	Location/Qualifiers
source	1. .175559
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	/clone="RP11-699A7"
repeat_region	1. .2494
	/rpt_family="L1M3f"
misc_feature	1. .2004
	/note="overlaps bases 123282. .125285 of clone AC121340"
	/function="clone overlap"
repeat_region	2492. .2613
	/rpt_family="L1"
repeat_region	2615. .3368
	/rpt_family="L1MA8"
repeat_region	3376. .3439
	/rpt_family="HAL1"
repeat_region	3540. .3926
	/rpt_family="HAL1"
repeat_region	3958. .4105
	/rpt_family="HAL1"
repeat_region	4469. .4494
	/rpt_family="(TTTA)n"
repeat_region	complement(4508. .4793)
	/rpt_family="AluY"
repeat_region	5704. .5739
	/rpt_family="(TTTG)n"
repeat_region	complement(5747. .6027)
	/rpt_family="AluJb"
repeat_region	6062. .6199
	/rpt_family="MLT1A2"
repeat_region	complement(6217. .6367)
	/rpt_family="L1MB2"
repeat_region	6854. .6888
	/rpt_family="(CAAG)n"
repeat_region	6891. .6929
	/rpt_family="GA-rich"
repeat_region	complement(7307. .7709)
	/rpt_family="LTR16A"
repeat_region	8153. .8249
	/rpt_family="MIR"
repeat_region	8323. .8343
	/rpt_family="AT_rich"
repeat_region	8344. .8528

```
                     /rpt_family="MIR"
     repeat_region   8536. .8573
                     /rpt_family="MIR"
     repeat_region   8789. .8973
                     /rpt_family="MER5A"
     repeat_region   9207. .9245
                     /rpt_family="L2"
     repeat_region   complement(10701. .10947)
                     /rpt_family="MLT1B"
     repeat_region   10948. .11189
                     /rpt_family="AluJb"
     repeat_region   complement(11190. .11344)
                     /rpt_family="MLT1B"
     repeat_region   complement(11500. .12669)
                     /rpt_family="L2"
     repeat_region   complement(12678. .12989)
                     /rpt_family="LTR16C"
     repeat_region   13075. .13115
                     /rpt_family="MADE1"
     repeat_region   13498. .13528
                     /rpt_family="AT_rich"
     repeat_region   complement(14373. .14598)
                     /rpt_family="MLT1L"
     repeat_region   15438. .15671
                     /rpt_family="MIR"
     repeat_region   complement(16593. .16897)
                     /rpt_family="MLT1L"
     repeat_region   17025. .17235
                     /rpt_family="L2"
     repeat_region   17868. .17980
                     /rpt_family="FLAM_C"
     misc_feature    17982. .20139
                     /note="Size of Repeat
                      2.15 KB; Verified by PCR and restriction digest"
                     /function="unresolved tandem repeat"
     repeat_region   17983. .18153
                     /rpt_family="(TA)n"
     repeat_region   complement(18163. .18193)
                     /rpt_family="HERVL"
     repeat_region   18242. .18362
                     /rpt_family="AluY"
     repeat_region   18398. .18560
                     /rpt_family="(TA)n"

  Query Match             22.8%;  Score 41.8;  DB 9;  Length 175559;
  Best Local Similarity   63.4%;  Pred. No. 0.0096;
  Matches   64;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
         |||||  | |  |||||| | |    | | | | ||||| |||||  ||   | ||| ||
Db 75968 GAGATTGAGGGACAGAAGAATAGCAAGATCATGATATTTGTTCCCTTAGCTACATCCATG 75909

Qy    112 CTGGGCTATGGATGAACAGTGGCTGACTTCATCTAGGAAAG 152
          | |||| |  | |   ||||| |||  ||| ||||| || |
Db  75908 CAGGGCCACTGGTTGGCAGTGTCTGCTTTCCTCTAGTAAGG 75868


RESULT 12
HS179D22
LOCUS       HS179D22              177220 bp    DNA     linear   PRI 24-JAN-2002
DEFINITION  Homo sapiens chromosome 9 BAC RP11-179D22, complete sequence.
ACCESSION   AL355433
VERSION     AL355433.5  GI:18376585
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 177220)
  AUTHORS   Scharfe,M., Conrad,A., Hornischer,K., Loehnert,T.H., Thies,S. and
            Bloecker,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2000) GBF, Dept. of Genome Analysis, Mascheroder
            Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
COMMENT     On Jan 25, 2002 this sequence version replaced gi:11140934.
            All annotations in this database entry are developed by
            computational tools. It is therefore not explicitly noted in the
            feature lines that evidence is not experimental.
            Mapping was performed at The Sanger Centre
            (cf. http://www.sanger.ac.uk/HGP/Chr9)
            Mapping information is available via
            http://webace.sanger.ac.uk/cgi-bin/display?db=acedb9&grep=179D22
            ------------- Genome Center
            Center: GBF, Braunschweig
            Center code: GBF
            Web site: http://genome.gbf.de/
            Contact: info.genome@gbf.de
            ------------- Project Information
            Center project name:
            Center clone name: bA179D22
            ------------- Summary Statistics
            Sequencing vector: ###;
            Chemistry: Dye-terminator-amersham: 85% of reads
            Chemistry: Dye-primer-amersham: 15% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 0 bases at least Q40
            Consensus quality: 0 bases at least Q30
            Consensus quality: 0 bases at least Q20
            Estimated insert size: ##; agarose-fp estimation
            Estimated insert size: 177220; sum-of-contigs estimation
            -------------.
            PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
            ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
            +   Analysis and annotation were performed with the automatic    +
            +          'first-pass' annotation and submission tool           +
            +             'AnnoMitter' (Hornischer & Bloecker).              +
            +                 Programs used by 'AnnoMitter':                 +
            ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
            > GeneFinder (Green), Vers. 084
            .   Organism: human
            > GenScan (Burge & Karlin), Vers. 1.0
            .   Used matrix: vertebrate; Minimum score: 0
            > Grail (Xu et al.), Vers. 1.3
            .   Organism: human
            > Mzef (Zhang)
            .   Prior probability: 0.04; Overlapping number: 0 > Xpound (Thomas
            & Skolnick)
            .   Base score cutoff: 0.2; Minimal exon length: 3 bp > 'Repeats':
            BLASTN 2.0.14 (Altschul et al.)
            .   Database(s): * RepBase: ALU (human), released 22-DEC-1995 .
                      * RepBase: THR ((human), released 22-DEC-1995 .
                  * RepBase: L1 (primate), released 22-DEC-1995 .
             * RepBase: MIR (primate), released 22-DEC-1995 .                *
            RepBase: MER (primate), released 22-DEC-1995 .                  *
            RepBase: MIR2 (primate), released 22-DEC-1995 .                 *
            RepBase: THE (primate), released 22-DEC-1995 .   Minimum score: 60;
            Minimum identity: 70 %;
            > 'ESTs': BLASTN 2.0.14 (Altschul et al.)
            .   Database(s): * embl (EST, human), released -DEC- .
              * embl (EST, other), released -DEC- .                 * emblnew
            (EST), Vers. 67+ (01-JAN-1970) .   Using sequence with masked
            repeats
            .   Minimum score: 60; Minimum identity: 70 %;
            > 'Tandem Repeats': GDE 2.2 option 'tandem'
            .   Minimum length 2 bp; Maximum length 20 bp; Score threshold 20 .
              Treat N's as mismatches? YES; Allow uniform consensi? NO >
            'Inverted Repeats': GDE 2.2 option 'inverted'
            > 'Micro Satellites': GDE 2.2 option 'sputnik' (Abajian) > 'CpG
            Islands': GDE 2.2 option 'cpg'
            .   CpG island region size 100 bp;
            .   Minimum GC contents 50 %; Observed/Expected 0.6 > 'STS Scan':
            e-PCR (Schuler)
            .   Margin: 50; Number of mismatches allowed: 0; Word size: 7 .
            STS database: 'dbSTS markers'
            > 'tRNA Scan': tRNAscan-SE (Lowe & Eddy), Vers. 1.11.
FEATURES              Location/Qualifiers
```

```
source          1. .177220
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="9"
                /clone="RP11-179D22"
                /clone_lib="RPCI-11.1"
misc_feature    1. .177220
                /note="assembly_fragment-clone_end:T7-vector_side:left
                assembly_fragment-clone_end:SP6-vector_side:right"
repeat_region   complement(4. .388)
                /note="84% identity: matches 1031. .1411 of consensus"
                /rpt_family="L1"
repeat_region   complement(389. .1013)
                /note="89% identity: matches 1373. .1996 of consensus"
                /rpt_family="L1"
exon            complement(486. .573)
                /note="GRAIL, score = 64%, comment = good"
repeat_region   complement(1010. .1100)
                /note="91% identity: matches 860. .950 of consensus"
                /rpt_family="L1"
repeat_region   complement(1116. .1500)
                /note="87% identity: matches 449. .836 of consensus"
                /rpt_family="L1"
exon            complement(1234. .1652)
                /note="GENSCAN prediction, score = 8.85"
exon            complement(1337. .1479)
                /note="GRAIL, score = 40%, comment = marginal"
repeat_region   complement(1496. .1689)
                /note="89% identity: matches 235. .428 of consensus"
                /rpt_family="L1"
repeat_region   complement(1692. .1915)
                /note="87% identity: matches 2. .219 of consensus"
                /rpt_family="L1"
repeat_region   complement(1942. .2269)
                /note="80% identity: matches 516. .830 of consensus"
                /rpt_family="L1"
exon            2044. .2078
                /note="XPOUND prediction, score = 0.241"
exon            complement(2179. .2291)
                /note="GRAIL, score = 54%, comment = good"
repeat_region   complement(2270. .2520)
                /note="86% identity: matches 252. .502 of consensus"
                /rpt_family="L1"
exon            complement(2333. .2469)
                /note="GENSCAN prediction, score = 5.49"
repeat_region   complement(2592. .2750)
                /note="87% identity: matches 6. .164 of consensus"
                /rpt_family="L1"
exon            complement(2678. .2728)
                /note="XPOUND prediction, score = 0.496"
exon            3019. .3074
                /note="MZEF prediction, score = 0.600"
exon            complement(3108. .3208)
                /note="XPOUND prediction, score = 0.736"
exon            complement(3112. .3208)
                /note="GENSCAN prediction, score = 14.35
                GRAIL, score = 100%, comment = excellent
                MZEF prediction, score = 0.536"
repeat_region   complement(3655. .3972)
                /note="86% identity: matches 673. .977 of consensus"
                /rpt_family="L1"
exon            complement(3767. .3779)
                /note="XPOUND prediction, score = 0.502"
repeat_region   complement(3955. .3992)
                /note="97% identity: matches 322. .359 of consensus"
                /rpt_family="L1"
exon            complement(4440. .4480)
                /note="GENSCAN prediction, score = 0.15"
exon            complement(5499. .5569)
                /note="GENSCAN prediction, score = 3.26
                GRAIL, score = 91%, comment = excellent
                MZEF prediction, score = 0.526"
exon            complement(7411. .7463)
                /note="XPOUND prediction, score = 0.357"
exon            8189. .8297
                /note="MZEF prediction, score = 0.521"
exon            complement(9660. .9762)
                /note="GENSCAN prediction, score = 5.86
                GRAIL, score = 73%, comment = good
                MZEF prediction, score = 0.925"
exon            complement(10293. .10550)
                /note="GRAIL, score = 100%, comment = excellent"
exon            complement(10307. .10550)
                /note="GENSCAN prediction, score = 26.53"
exon            complement(10308. .10596)
                /note="XPOUND prediction, score = 0.819"
misc_feature    10388. .10525
                /note="CpG_island (%GC=62.3, o/e=0.73, #CpGs=11)"
satellite       11243. .11253
                /note="TC repeat"
repeat_region   11499. .11583
                /note="84% identity: matches 12. .96 of consensus"
                /rpt_family="ALU"
satellite       11615. .11628
                /note="AAAAT repeat"
repeat_region   11630. .11676
                /note="91% identity: matches 133. .179 of consensus"
                /rpt_family="ALU"
repeat_region   11720. .11778
                /note="91% identity: matches 216. .274 of consensus"
                /rpt_family="ALU"
satellite       11780. .11801
                /note="CTAT repeat"
repeat_region   11959. .12128
                /note="homology = 74.1%, counts = 85"
                /rpt_family="at repeat"
                /rpt_type=TANDEM
satellite       12073. .12085
                /note="AT repeat"
satellite       12115. .12128
                /note="AT repeat"
repeat_region   complement(12142. .12412)
                /note="89% identity: matches 8. .273 of consensus"

  Query Match             22.8%;  Score 41.8;  DB 9;  Length 177220;
  Best Local Similarity   75.4%;  Pred. No. 0.0096;
  Matches   52;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      52 GAGATCCAAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTG 111
           || || |||| |  |||||| | |||||| ||||||||||||  |||||  || || |||
Db   79922 GACATACAAGGGTGGAAGGACAGTGTGGTCAGGCTATTTATTTTCCCAGCTCCGTCTCTG 79981

Qy     112 CTGGGCTAT 120
             || || ||
Db   79982 TTGAGCCAT 79990


RESULT 13
AC107875
LOCUS       AC107875               58693 bp    DNA     linear   HTG 13-FEB-2002
DEFINITION  Homo sapiens chromosome 8 clone RP11-640C18 map 8, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC107875
VERSION     AC107875.2  GI:18653637
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58693)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-640C18
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 58693)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
```

Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Feb 13, 2002 this sequence version replaced gi:18308592.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L22007
Center clone name: 640_C_18
----------.
* NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 679: contig of 679 bp in length
* 680 779: gap of 100 bp
* 780 1492: contig of 713 bp in length
* 1493 1592: gap of 100 bp
* 1593 2305: contig of 713 bp in length
* 2306 2405: gap of 100 bp
* 2406 3042: contig of 637 bp in length
* 3043 3142: gap of 100 bp
* 3143 3855: contig of 713 bp in length
* 3856 3955: gap of 100 bp
* 3956 4648: contig of 693 bp in length
* 4649 4748: gap of 100 bp
* 4749 5465: contig of 717 bp in length
* 5466 5565: gap of 100 bp
* 5566 6259: contig of 694 bp in length
* 6260 6359: gap of 100 bp
* 6360 7056: contig of 697 bp in length
* 7057 7156: gap of 100 bp
* 7157 7810: contig of 654 bp in length
* 7811 7910: gap of 100 bp
* 7911 8627: contig of 717 bp in length
* 8628 8727: gap of 100 bp
* 8728 9400: contig of 673 bp in length
* 9401 9500: gap of 100 bp
* 9501 10214: contig of 714 bp in length
* 10215 10314: gap of 100 bp
* 10315 11020: contig of 706 bp in length
* 11021 11120: gap of 100 bp
* 11121 11846: contig of 726 bp in length
* 11847 11946: gap of 100 bp
* 11947 12661: contig of 715 bp in length
* 12662 12761: gap of 100 bp
* 12762 13467: contig of 706 bp in length
* 13468 13567: gap of 100 bp
* 13568 14271: contig of 704 bp in length
* 14272 14371: gap of 100 bp
* 14372 15069: contig of 698 bp in length
* 15070 15169: gap of 100 bp
* 15170 15882: contig of 713 bp in length
* 15883 15982: gap of 100 bp
* 15983 16700: contig of 718 bp in length
* 16701 16800: gap of 100 bp
* 16801 17501: contig of 701 bp in length
* 17502 17601: gap of 100 bp
* 17602 18328: contig of 727 bp in length
* 18329 18428: gap of 100 bp
* 18429 19117: contig of 689 bp in length
* 19118 19217: gap of 100 bp
* 19218 19925: contig of 708 bp in length
* 19926 20025: gap of 100 bp
* 20026 20755: contig of 730 bp in length
* 20756 20855: gap of 100 bp
* 20856 21566: contig of 711 bp in length
* 21567 21666: gap of 100 bp
* 21667 22374: contig of 708 bp in length
* 22375 22474: gap of 100 bp
* 22475 23181: contig of 707 bp in length
* 23182 23281: gap of 100 bp
* 23282 24001: contig of 720 bp in length
* 24002 24101: gap of 100 bp
* 24102 24811: contig of 710 bp in length
* 24812 24911: gap of 100 bp
* 24912 25600: contig of 689 bp in length
* 25601 25700: gap of 100 bp
* 25701 26422: contig of 722 bp in length
* 26423 26522: gap of 100 bp
* 26523 27228: contig of 706 bp in length
* 27229 27328: gap of 100 bp
* 27329 28041: contig of 713 bp in length
* 28042 28141: gap of 100 bp
* 28142 28844: contig of 703 bp in length
* 28845 28944: gap of 100 bp
* 28945 29628: contig of 684 bp in length
* 29629 29728: gap of 100 bp
* 29729 30423: contig of 695 bp in length
* 30424 30523: gap of 100 bp
* 30524 31235: contig of 712 bp in length
* 31236 31335: gap of 100 bp
* 31336 31979: contig of 644 bp in length
* 31980 32079: gap of 100 bp
* 32080 32797: contig of 718 bp in length
* 32798 32897: gap of 100 bp
* 32898 33595: contig of 698 bp in length
* 33596 33695: gap of 100 bp
* 33696 34416: contig of 721 bp in length
* 34417 34516: gap of 100 bp
* 34517 35235: contig of 719 bp in length
* 35236 35335: gap of 100 bp
* 35336 36034: contig of 699 bp in length
* 36035 36134: gap of 100 bp
* 36135 36846: contig of 712 bp in length
* 36847 36946: gap of 100 bp
* 36947 37670: contig of 724 bp in length
* 37671 37770: gap of 100 bp
* 37771 38479: contig of 709 bp in length
* 38480 38579: gap of 100 bp
* 38580 39287: contig of 708 bp in length
* 39288 39387: gap of 100 bp
* 39388 40105: contig of 718 bp in length

```
          *    40106 40205: gap of      100 bp
          *    40206    40918: contig of 713 bp in length
          *    40919 41018: gap of      100 bp
          *    41019    41723: contig of 705 bp in length
          *    41724 41823: gap of      100 bp
          *    41824    42495: contig of 672 bp in length
          *    42496 42595: gap of      100 bp
          *    42596    43285: contig of 690 bp in length
          *    43286 43385: gap of      100 bp
          *    43386    44095: contig of 710 bp in length
          *    44096 44195: gap of      100 bp
          *    44196    44894: contig of 699 bp in length
          *    44895 44994: gap of      100 bp
          *    44995    45705: contig of 711 bp in length
          *    45706 45805: gap of      100 bp
          *    45806    46517: contig of 712 bp in length
          *    46518 46617: gap of      100 bp
          *    46618    47326: contig of 709 bp in length
          *    47327 47426: gap of      100 bp
          *    47427    48155: contig of 729 bp in length
          *    48156 48255: gap of      100 bp
          *    48256    48985: contig of 730 bp in length
          *    48986 49085: gap of      100 bp
          *    49086    49803: contig of 718 bp in length
          *    49804 49903: gap of      100 bp
          *    49904    50623: contig of 720 bp in length
          *    50624 50723: gap of      100 bp
          *    50724    51432: contig of 709 bp in length
          *    51433 51532: gap of      100 bp
          *    51533    52248: contig of 716 bp in length
          *    52249 52348: gap of      100 bp
          *    52349    53068: contig of 720 bp in length
          *    53069 53168: gap of      100 bp
          *    53169    53881: contig of 713 bp in length
          *    53882 53981: gap of      100 bp
          *    53982    54683: contig of 702 bp in length
          *    54684 54783: gap of      100 bp

  Query Match             22.7%;  Score 41.6;  DB 2;  Length 58693;
  Best Local Similarity   62.5%;  Pred. No. 0.011;
  Matches   65;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy      59 AAGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCT 118
           | | |||| |||||| |||||| | | | || ||| |    |   | |||||  || ||
Db   38355 AGGGGCAGGAGGAGAGTGTGGTCAAGGTCTTCATTTCTTTTGCTACCTCCCTATTGTGCC 38414

Qy     119 ATGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGCTATGGCT 162
           ||| || | ||||| ||   ||  ||||  |||  | ||   ||
Db   38415 ATGAATTATCAGTGTCTTTGTTTCTCTACCAAAAGGATAGCTCT 38458


RESULT 14
AP005368/c
LOCUS       AP005368              111461 bp    DNA     linear   PRI 30-JUL-2002
DEFINITION  Homo sapiens genomic DNA, chromosome 8q23, clone: KB870F5, complete
            sequence.
ACCESSION   AP005368
VERSION     AP005368.2  GI:22004069
KEYWORDS    .
SOURCE      Homo sapiens pre-pro-B cell cell_line:FLEB 14 - 14 DNA,
            clone_lib:Keio BAC library clone:KB870F5.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimizu,N. and Asakawa,S.
  TITLE     Homo sapiens DNA chromosome 8 SEQUENCE
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 111461)
  AUTHORS   Shimizu,N. and Asakawa,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-2002) Nobuyoshi Shimizu, Keio University, School
            of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo
            160-8582, Japan (E-mail:nshimizu@dmb.med.keio.ac.jp,
            Tel:81-3-3351-2370, Fax:81-3-3351-2370)
COMMENT     On Jul 30, 2002 this sequence version replaced gi:21280421.
FEATURES             Location/Qualifiers
     source          1. .111461
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q23"
                     /clone="KB870F5"
                     /cell_line="FLEB 14 - 14"
                     /cell_type="pre-pro-B cell"
                     /clone_lib="Keio BAC library"
     repeat_region   1. .5368
                     /evidence=not_experimental
                     /rpt_family="L1PA2"
     repeat_region   5369. .5399
                     /evidence=not_experimental
                     /rpt_family="(TAA)n"
     repeat_region   5408. .5787
                     /evidence=not_experimental
                     /rpt_family="L1MD"
     repeat_region   5828. .5920
                     /evidence=not_experimental
                     /rpt_family="L1MD"
     repeat_region   6023. .7911
                     /evidence=not_experimental
                     /rpt_family="L1MEc"
     repeat_region   8036. .8194
                     /evidence=not_experimental
                     /rpt_family="L1PA16"
     repeat_region   8415. .8537
                     /evidence=not_experimental
                     /rpt_family="LTR47A"
     repeat_region   8541. .9383
                     /evidence=not_experimental
                     /rpt_family="L1M4"
     repeat_region   9420. .9550
                     /evidence=not_experimental
                     /rpt_family="L1MD2"
     repeat_region   complement(9572. .9750)
                     /evidence=not_experimental
                     /rpt_family="L2"
     repeat_region   complement(10365. .10571)
                     /evidence=not_experimental
                     /rpt_family="MER20"
     repeat_region   10574. .10743
                     /evidence=not_experimental
                     /rpt_family="AluSg/x"
     repeat_region   complement(10779. .10898)
                     /evidence=not_experimental
                     /rpt_family="MIR"
     repeat_region   complement(11064. .11635)
                     /evidence=not_experimental
                     /rpt_family="MLT1E1"
     repeat_region   11795. .12089
                     /evidence=not_experimental
                     /rpt_family="AluSc"
     repeat_region   complement(12308. .12499)
                     /evidence=not_experimental
                     /rpt_family="MIR"
     repeat_region   complement(12741. .13050)
                     /evidence=not_experimental
                     /rpt_family="AluSg"
     repeat_region   complement(13538. .13866)
                     /evidence=not_experimental
                     /rpt_family="L2"
     repeat_region   complement(13903. .14303)
                     /evidence=not_experimental
                     /rpt_family="L2"
     repeat_region   complement(15438. .15626)
                     /evidence=not_experimental
                     /rpt_family="MIR"
```

```
repeat_region   complement(15961. .16152)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(17363. .17408)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   17596. .17901
                /evidence=not_experimental
                /rpt_family="AluY"
repeat_region   17995. .18317
                /evidence=not_experimental
                /rpt_family="AluJo"
repeat_region   18840. .24954
                /evidence=not_experimental
                /rpt_family="L1PA2"
repeat_region   24955. .24981
                /evidence=not_experimental
                /rpt_family="(A)n"
repeat_region   complement(25025. .25248)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   25426. .25473
                /evidence=not_experimental
                /rpt_family="(CA)n"
repeat_region   25804. .25959
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   26581. .26744
                /evidence=not_experimental
                /rpt_family="MER5B"
repeat_region   28119. .29663
                /evidence=not_experimental
                /rpt_family="L1MD2"
repeat_region   29664. .29924
                /evidence=not_experimental
                /rpt_family="AluSg"
repeat_region   29925. .32074
                /evidence=not_experimental
                /rpt_family="L1MD2"
repeat_region   32671. .33096
                /evidence=not_experimental
                /rpt_family="L1ME3A"
repeat_region   complement(33097. .33578)
                /evidence=not_experimental
                /rpt_family="MLT1D"
repeat_region   33579. .33988
                /evidence=not_experimental
                /rpt_family="L1ME3A"
repeat_region   34067. .34230
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   35148. .35425
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   complement(36218. .36247)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(36306. .36770)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   complement(37600. .37656)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   complement(37784. .38171)
                /evidence=not_experimental
                /rpt_family="MSTA"
repeat_region   39215. .39545
                /evidence=not_experimental
                /rpt_family="MLT1G1"
repeat_region   39716. .39802
                /evidence=not_experimental
                /rpt_family="L1MA8"
repeat_region   complement(39803. .40124)
                /evidence=not_experimental
                /rpt_family="MER2"
repeat_region   40125. .40294
                /evidence=not_experimental
                /rpt_family="L1MA8"
repeat_region   complement(40415. .40456)
                /evidence=not_experimental
                /rpt_family="5S"
repeat_region   complement(40667. .40900)
                /evidence=not_experimental
                /rpt_family="L1M4"
repeat_region   43939. .45511
                /evidence=not_experimental
                /rpt_family="L1MEc"
repeat_region   complement(45600. .45863)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   complement(45868. .45935)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region   45937. .46431
                /evidence=not_experimental
                /rpt_family="MLT1D"
repeat_region   complement(47468. .47673)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   47724. .47750
                /evidence=not_experimental
                /rpt_family="(TC)n"
repeat_region   48131. .48249
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region   complement(48250. .48572)
                /evidence=not_experimental

  Query Match             22.7%;  Score 41.6;  DB 9;  Length 111461;
  Best Local Similarity   64.6%;  Pred. No. 0.011;
  Matches   62;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      60 AGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTA 119
           || |||| |||||| |||||| | | | || ||| |    |   | |||||  || || |
Db   90636 AGGGCAGGAGGAGAGTGTGGTCAAGGTCTTCATTTCTTTTGCTACCTCCCTATTGTGCCA 90577

Qy     120 TGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGC 155
           || || | ||||| ||   ||  ||||  |||  ||
Db   90576 TGAATTATCAGTGTCTTTGTTTCTCTACCAAAAGGC 90541


RESULT 15
AC105177/c
LOCUS       AC105177               149126 bp    DNA     linear   PRI 01-MAY-2002
DEFINITION  Homo sapiens chromosome 8, clone RP11-91D24, complete sequence.
ACCESSION   AC105177
VERSION     AC105177.5  GI:20336821
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 149126)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-91D24
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 149126)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
```

Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (26-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 149126)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (29-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 149126)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (01-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Apr 29, 2002 this sequence version replaced gi:20334653.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L21920
Center clone name: 91_D_24
----------.

FEATURES	Location/Qualifiers
source	1. .149126
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="8"
	/map="8"
	/clone="RP11-91D24"
	/clone_lib="RPCI-11 Human Male BAC"
repeat_region	complement(205. .383)
	/rpt_family="L2"
repeat_region	complement(998. .1204)
	/rpt_family="MER20"
repeat_region	1207. .1376
	/rpt_family="AluSg/x"
repeat_region	complement(1412. .1548)
	/rpt_family="MIR"
repeat_region	complement(1697. .2305)
	/rpt_family="MLT1E1"
repeat_region	2428. .2722
	/rpt_family="AluSc"
unsure	complement(2800. .2824)
	/note="<30 SNGL Qual Region"
unsure	complement(2830. .2837)
	/note="< 30 Qual SNGL Region"
unsure	complement(2851. .2855)
	/note="< 30 Qual SNGL Region"
unsure	complement(2863. .2874)
	/note="<30 Qual SNGL Region"
unsure	complement(2895. .2901)
	/note="<30 qual SNGL Region"
repeat_region	complement(2938. .3129)
	/rpt_family="MIR"
repeat_region	complement(3371. .3680)
	/rpt_family="AluSg"
repeat_region	complement(4168. .4496)
	/rpt_family="L2"
repeat_region	complement(4643. .4971)
	/rpt_family="L2"
repeat_region	complement(6068. .6256)
	/rpt_family="MIR"
repeat_region	complement(6591. .6782)
	/rpt_family="MIR"
repeat_region	complement(7868. .7972)
	/rpt_family="MIR3"
repeat_region	complement(7993. .8038)
	/rpt_family="L2"
repeat_region	8226. .8531
	/rpt_family="AluY"
repeat_region	8625. .8948
	/rpt_family="AluSx"
repeat_region	complement(9153. .9225)
	/rpt_family="MIR"
repeat_region	9471. .15592
	/rpt_family="L1PA5"
repeat_region	complement(15655. .15878)
	/rpt_family="MIR"
repeat_region	16056. .16103
	/rpt_family="(CA)n"
repeat_region	16434. .16589
	/rpt_family="MIR"
repeat_region	17211. .17374
	/rpt_family="MER5B"
repeat_region	18749. .20293
	/rpt_family="L1MD2"
repeat_region	20294. .20554

```
                     /rpt_family="AluSg"
     repeat_region   20555. .22704
                     /rpt_family="L1MD2"
     repeat_region   23301. .23726
                     /rpt_family="L1ME3A"
     repeat_region   complement(23727. .24208)
                     /rpt_family="MLT1D"
     repeat_region   24209. .24618
                     /rpt_family="L1ME3A"
     repeat_region   24697. .24860
                     /rpt_family="MIR"
     repeat_region   25926. .26055
                     /rpt_family="L2"
     repeat_region   complement(26936. .27397)
                     /rpt_family="L2"
     repeat_region   complement(28230. .28286)
                     /rpt_family="L2"
     repeat_region   complement(28414. .28801)
                     /rpt_family="MSTA"
     repeat_region   29845. .30164
                     /rpt_family="MLT1G1"
     repeat_region   30346. .30432
                     /rpt_family="L1MA8"
     repeat_region   complement(30433. .30755)
                     /rpt_family="MER2"
     repeat_region   30756. .30925
                     /rpt_family="L1MA8"
     repeat_region   complement(31132. .31498)
                     /rpt_family="L1M4"
     repeat_region   34570. .36142
                     /rpt_family="L1MEc"

  Query Match             22.7%;  Score 41.6;  DB 9;  Length 149126;
  Best Local Similarity   64.6%;  Pred. No. 0.011;
  Matches   62;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      60 AGTGCAGAAGGAGAATGTGGTGAGGCTATTTATTCCCCCAGTGCCTTCCCTGCTGGGCTA 119
           || |||| |||||| |||||| | | | || ||| |    |   | |||||  || || |
Db   81299 AGGGCAGGAGGAGAGTGTGGTCAAGGTCTTCATTTCTTTTGCTACCTCCCTATTGTGCCA 81240

Qy     120 TGGATGAACAGTGGCTGACTTCATCTAGGAAAGAGC 155
           || || | ||||| ||   ||  ||||  |||  ||
Db   81239 TGAATTATCAGTGTCTTTGTTTCTCTACCAAAAGGC 81204


Search completed: May 25, 2003, 03:59:04
Job time : 2706 secs
```